UNIVERSITY OF CALIFORNIA, SAN DIEGO

Financial Volatility and the Macroeconomy

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Economics

by

Sung Je Byun

Committee in charge:

Professor James D. Hamilton, Chair
Professor Valerie A. Ramey
Professor Allan Timmermann
Professor Alexis A. Toda
Professor Ruth J. Williams

2015

The dissertation of Sung Je Byun is approved, and
it is acceptable in quality and form for publication
on microfilm:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2015

DEDICATION

To my father, Young J. Byun, who is still living in me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

It has been a long road, but I have come through the other end. When my dad alluded to his hope in 2003, I promised to earn a Ph.D naively, without knowing how many additional years that I would stay in college. It was the beginning of my long journey of scholarship and went smoothly until the sudden death of my father in 2006, in which I returned to Seoul in order to look after the bereaved mother. Though not an academic career, my job as a mutual fund manager at Mirae Asset Global Investments was good enough to make me her proud son. However, I decided to resume my journey to be a scholar not only for the unfulfilled desire of my late father, but also to fulfill a life's calling. While studying at University of California, San Diego since September 2009, I would never have been able to finish my journey without the guidance of my committee members, assistance from friends, and support from my family and wife.

I would like to express my deepest gratitude to my advisor, Professor James Hamilton, for his excellent guidance, care and patience. Although challenged to satisfy his high standard of discipline, I learned the A to Z of research from Jim, which ranges from formulating research questions into appropriate models/methods to presenting results effectively in public. Thanks to his careful concern and patience, not only did I complete my Ph.D, but I was also able to develop my writing skills as a powerful tool for sharing my research ideas effectively.

I would also like to dedicate my gratitude to my other committee members, Professor Valerie Ramey, Professor Allan Timmermann and Professor Ruth Williams, for guiding my research for the past several years. Special thanks go to Professor Alexis Toda, who was willing to participate in my committee at the last moment and provided helpful comments during the job search process, and also to Professor Thomas Baranga, who was a good professor, a thoughtful employer and has become a fabulous friend while in my graduate years at UCSD.

I would like to thank Martha Stacklin, who was willing to help my preparations for all presentations as well as acted as a matchmaker to discover my wife's

inner beauty. Undoubtedly I would not have survived at UCSD without her endless care. I am also deeply grateful to Soojin Jo, who had spent countless hours sharing and discussing research questions with me. Many thanks to all of my colleagues - current and former - for helping me survive all the stress from this year and not letting me give up.

I would like to express my special thanks to the two most important women in my life, whom I love beyond all words: my mother Soon-Ok Lee, who taught me the wisdom of patience, and my wife Guipeun Kang, who has fully supported my journey, despite her own challenges in seeking for a Ph.D in Chemistry. Finally, I am sincerely grateful to my siblings, Sung Hoon and Soo Hyun, for their great faith, and also to my parents-in-law for their endless support and prayers.

Chapters 1 and 2, in full, have been submitted for publication of material. Chapter 3, "Heterogeneity in the Dynamic Effects of Uncertainty on Investment", is coauthored with Soojin Jo, and in full, has been submitted for publication of material. I thank Soojin for the permission to use this chapter in my dissertation.

<div align="center">VITA</div>

2004	B. A. in Business Administration and B. A. in Applied Statistics, Yonsei University, Seoul, Korea
2006	M. A in Statistics, Columbia University, New York
2007-2009	Associate Portfolio Manager, Mirae Asset Global Investments, Seoul, Korea
2015	Ph. D. in Economics, University of California, San Diego

ABSTRACT OF THE DISSERTATION

Financial Volatility and the Macroeconomy

by

Sung Je Byun

Doctor of Philosophy in Economics

University of California, San Diego, 2015

Professor James D. Hamilton, Chair

This dissertation studies the effect of financial asset volatility on the macroeconomy. As an important source of information, I use cross-sectional dispersion for improving volatility forecasts along with time-variation in financial assets.

The first chapter, "Speculation in Commodity Futures Market, Inventories and the Price of Crude Oil", investigates the effects of financial investors activities in commodity markets on crude oil price. While earlier researchers addressed this question based on proxies representing financial investors activities, I develop a model of the convenience yield arising from holding crude oil inventories in spite of anticipated falling prices. Although some have argued that a breakdown of the relationship between crude oil inventories and prices following increased participation by financial investors after 2003 was evidence of an effect of speculation, I find that a correctly specified relation is stable over time. In light of this new evidence, I conclude that the contribution of financial investors activities is weak in the crude oil market.

In the second chapter, "The Usefulness of Cross-sectional Dispersion for

Forecasting Aggregate Stock Price Volatility", I develop a model of stock returns where dispersion in returns across different stocks is modeled jointly with aggregate volatility. Although specifications that allow for feedback from cross-sectional dispersion to aggregate volatility have a better fit in sample, I find that such improvements are not robust for purposes of out-of-sample forecasting. Using a full cross-section of stock returns jointly, however, I find that use of cross-sectional dispersion can help improve estimates of the parameters of a GARCH process for aggregate volatility, providing better forecasts both in sample and out of sample. Given this evidence, I conclude that cross-sectional information helps predict market volatility indirectly rather than directly entering in the data-generating process.

My final dissertation chapter, "Heterogeneity in the Dynamic Effects of Uncertainty on Investment", studies the effects of profit uncertainty on manufacturing firms investment decisions. We measure aggregate profit uncertainty from quarterly industry-level sales revenues by using a Panel-ARCH model, which is a special case of the bivariate aggregate volatility model developed in the second chapter. Using estimated profit uncertainty, we find that higher profit uncertainty induces firms to lower future capital expenditure on average, yet to a different degree depending on each firms characteristics, such as size, liability ratio, and sub-industry classification. This finding points to significant and substantial heterogeneity in the uncertainty transmission mechanism, a feature not highlighted in recent studies of uncertainty at the aggregate level.

Chapter 1

Speculation in Commodity Futures Markets, Inventories and the Price of Crude Oil

Abstract. Refiners have a motive to hold inventories even if they antici-
pate falling crude oil prices. This paper develops a model of the convenience yield
arising from holding inventories. Although some have argued that a breakdown of
the relationship between crude oil inventories and prices following increased par-
ticipation by financial investors after 2003 was evidence of an effect of speculation,
we find that a correctly specified relationship is stable over time. In light of this
evidence, we conclude that the contribution of financial investors' activities is weak
in the crude oil market.

1.1 Introduction

The recent volatility of crude oil prices has renewed interest in the behavior of crude oil inventories. This paper examines the role of inventories in refiners' gasoline production, and develops a structural model of the relation between crude oil prices and inventories.

In a competitive commodity market, a producer makes the optimal storage decision by equating his expected benefit with the relevant cost of holding inventories. The positive benefit could motivate a producer to hold inventories even if he anticipates falling prices in the future. Given this motivation, the importance of inventories for storable commodities has been widely recognized in the theory of storage literature. Among earlier researchers, [17] defined this benefit as the convenience yield, which is the flow of services that accrues to an owner of the physical commodity, but not to an owner of a contract for future delivery of the commodity. Most earlier literature focused on the inventories' role of reducing future production cost and empirically tested the convenience yield in various storable commodities.

In this paper, we propose an equilibrium storage model of the global crude oil market to study the role of inventories in refiners' gasoline production and to explain fluctuations of crude oil prices in terms of refiners' benefits to hold inventories. First, we model determinants of crude oil inventories, building upon earlier research by [39] and [30]. Specifically, we model the role of crude oil inventories directly as enhancing refiners' gasoline production by treating inventories as an essential factor of production following Kydland and Prescott [76, 1988] and [90]. Second, we show the convenience yield is inversely related to crude oil inventories, yet positively related to crude oil prices. This finding is consistent with the earlier conjecture in the theory of storage. Furthermore, our model estimates indicate that the convenience yield is higher for the summer and for the fall than for other seasons, which is consistent with observed seasonality in crude oil inventories.

Third, we identify the structural change in crude oil market fundamentals since 2004 from a rising permanent component of crude oil prices. It is also evidenced in changing model parameter estimates for before/after 2004, and is consistent with earlier literature (Tang and Xiong 2012; Büyükşahin and Robe 2014; Hamilton and Wu 2014).

We illustrate one use of our framework by re-examining the possible contribution of financial investors using a proposed storage model. Given volatile commodity prices and rapidly increasing financial investors' activities since 2004, it has been a controversial issue whether the increasing financial investors' participation has influenced commodity prices in recent years. However, empirical evidence among researchers is inconsistent. For example, some studies find evidence that financial investors' activities have impacted commodity futures prices (See, e.g., Büyüksahin et al. 2008; Einloth 2009; Gilbert 2010; Tang and Xiong 2012; Singleton 2013; Büyükşahin and Robe 2014). On the other hand, others find little evidence of a relationship between financial investors' activities and movements in commodity futures prices (See, e.g., Büyüksahin et al. 2008; Stoll and Whaley 2010; Irwin and Sanders 2010; Sanders and Irwin 2010; Aulerich et al. 2010; Büyüksahin and Harris 2011; Fattouh and Mahadeva 2012; Kilian and Murphy 2014).

One piece of evidence that financial investors have influenced crude oil prices was provided by [80]. They noted that traditional models of the relation between crude oil prices and inventories, such as that developed by [105], broke down after financial investor participation in crude oil markets increased. Merino and Ortiz extended these models to include a contribution of the long positions of non-commercial traders, and found that such a specification could better explain the relation between prices and inventories over 2001-2004. In this paper, we reproduce the good in-sample fit of the Ye specification, its breakdown after 2001, and the improvement provided by non-commercial positions over 2001-2004. However, we also find that even the latter relation falls apart after 2004. We show that although

the traditional model seemed to do well over the period 1992-2001, it was in fact misspecified even over that period, and simply appeared to do well because there was relatively little change in the permanent component of crude oil prices over that period. We show that the model of the inventory-price relation proposed here is stable over time and furthermore can account for both the apparent success of the [105] specification on earlier data as well as its breakdown on subsequent data.

The remainder of the paper proceeds as follows: In section 1.2, we introduce the theoretical model where the convenience yield arises from the representative producer's expected future productivity gain from holding crude oil inventories. After introducing ways to deal with three issues such as the crude oil spot price, the aggregation and the seasonality, the equilibrium prediction is provided at the end of section 1.2. In section 1.3, we begin with documenting estimation results together with explanations for MLE estimates of structural model parameters. Next, we show the strong forecasting relationship of inventories and prices, indicating the weak contribution of financial investors' activities in the crude oil market. After providing robustness checks, we explore the recent episode in the crude oil market based on the stable inventory-price relationship in section 1.4. Lastly, we conclude with some remarks.

1.2   Theoretical Model

In this section, we begin by introducing the accounting identity of inventories and the cost function of the representative producer. Next, we provide the role of inventories as facilitating production schedule and avoiding stockouts by treating inventories as factors of production. Combining the cost and the production functions, a theoretical model is provided in the context of the representative producer's decision problem. Lastly, after dealing with issues related to the formulation of the crude oil spot price process, the aggregation and the seasonality, we yield the equilibrium prediction at the aggregated global economy level.

1.2.1   Cost Function

Consider a representative producer who purchases a quantity q_t^P of crude oil at price S_t per barrel, of which q_t^U is used up in current production of consumption good such as gasoline and the remainder goes to increase inventories[1] (i_t):

$$i_t = i_{t-1} + q_t^P - q_t^U, \tag{1.1}$$

where $q_t^P - q_t^U$ corresponds to net additional crude oil inventories during the time period t. If the quantity of oil consumed by the representative producer (q_t^U) is smaller than the quantity purchased (q_t^P), inventories accumulate and vice versa.

The cost function summarizes both the production and storage technology for the representative producer. Given the current crude oil spot price (S_t), the firm's costs come from two components: costs for purchasing resources and those for storing inventories over one period. The former is the product of crude oil spot price and amounts purchased (q_t^P), and the latter is assumed to be proportional to the current crude oil spot price[2]. For reflecting the idea of limited storage facilities

---

[1]Most earlier literature focused on finished good inventories in order to explain roles such as production smoothing and production cost smoothing. In what follows, we focus on raw material inventories and explain the benefit from holding raw material inventories.

[2]We define the cost for storage in a broad sense: it includes the cost for insurance and

in the short-run, storage costs are further assumed to take the form of a convex quadratic function. Given the current crude oil spot price (S_t) and previous level of inventories (i_{t-1}), the representative producer's cost function becomes,

$$c\left(q_t^U, i_t; i_{t-1}, S_t\right) = S_t \left\{q_t^U + (i_t - i_{t-1})\right\} + S_t \cdot \left(c_0 + \frac{c_1}{2}i_t^2\right),$$

where we plug in the accounting identity of inventories following (1.1) after rearranging it for a quantity purchased (q_t^P) and $c_0, c_1 > 0$.

1.2.2 Production Function

The importance of inventories has been widely recognized in the theory of storage developed by [72], [104] and [17]. In particular, [17] defined the the convenience yield as the flow of services that accrues to an owner of the physical commodity but not to an owner of a contract for future delivery of the commodity. These services can arise from reducing the probability of a stock-out of inventories, from inventories' role in production smoothing, and from future production cost saving. For example, [84] described the role of inventories in production is "to reduce costs of adjusting production and to reduce marketing costs by facilitating schedule and avoiding stockouts". With his proposed measure for the intangible convenience yield using available futures contract prices, he found evidence for such roles for copper, heating oil, and lumber. Further empricial tests for the convenience yield in commodities include Fama and French [46, 1988], [32], [81], [91], [19], [59], [3], [87].

The role of inventories as a factor of production has been widely adopted in earlier literature. For example, Kydland and Prescott [76, 1988] and [90] treat inventories as essential factors of production for studies of aggregate fluctuations in the U.S. economy. Focusing on the petroleum refining industry, [30] examines determinants of inventory investment under joint production by treating inventories

---

transportation beside that for using storage facilities. In general, these costs are determined as being proportional to the value that is stored.

as quasi-fixed factors of production. In this paper, we model crude oil inventories as directly increasing the refinery's production capabilities. In order to motivate this role of inventories further, consider the refinery's decision problem for a periodic production schedule under resource constraints. Compared to firms in other industries, the refinery's production heavily depends on the raw material, in this case, crude oil. It generally takes more than a few weeks to receive additional crude oil delivery after placing an order in the spot market or making transaction in the derivative market. For its current gasoline production, the refinery can use at most crude oil resources that are either carried over from the previous period or are delivered currently from the past transaction. In response to uncertainties in the aggregate gasoline demand and petroleum prices, the available inventories in the refinery's storage determine the attainable level of production and can improve production efficiency without needing to make costly production adjustments. Hence, the refinery has a motivation for holding inventories despite an anticipated falling crude oil prices in the future. Given the realized technology process (z_t), we propose a production function of the form,

$$f\left(q_t^U; i_{t-1}, z_t\right) = \left(e^{z_t} q_t^U\right)^\alpha \left\{1 - \exp\left(-\theta_t \cdot i_{t-1}\right)\right\},$$

where $\alpha \in (0,1)$ is the output share of the crude oil resource and $\theta_t > 0$ is the utilization parameter governing the production function's dependence on the previous level of inventory. The subscript t on utilization parameter (θ_t) is used to allow the possiblity for taking different values depending on the season of the year, for example, taking a smaller value when the refinery tries to produce more gasoline for the summer driving season. The term in curly brackets is bounded between 0 and 1, approaching 1 for sufficiently large i_{t-1} and approaching 0 for i_{t-1} being close to 0. The resource-augmenting technology process (z_t) follows the random walk process, that is, $z_{t+1} = z_t + \epsilon_{1,t+1}$ with $\epsilon_{1,t+1} \sim N\left(0, \sigma_1^2\right)$.

1.2.3 Theoretical Model

With the cost function and the production function as introduced earlier, the representative producer faces a dynamic programming problem. At the beginning of each period t, the representative producer faces the crude oil spot price of S_t, the realized technology process z_t and the exogenously determined real interest rate r_t. Given previous inventory (i_{t-1}), he makes decisions on the resource demand (q_t^U) and inventories (i_t) jointly in order to maximize his lifetime profits. Suppose he is a price taker in the crude oil market and he is risk averse with discount factor (Λ_t),[3] which governs his risk-aversion. Suppressing the superscript in the resource demand (q_t), the representative producer's objective is thus to choose $\{q_t, i_t\}_{t=1}^{\infty}$ so as to maximize

$$\Pi = \max_{\{q_t, i_t\}} E_0 \left[\sum_{t=1}^{\infty} \prod_{\tau=1}^{t} \Lambda_{\tau} e^{-r_{\tau}} \left\{ f\left(q_t; i_{t-1}, z_t\right) - c\left(q_t, i_t; i_{t-1}, S_t\right) \right\} \right],$$

where i_{-1} is given. Note we pose this as the representative agent's problem of producing consumption goods rather than gasoline, for the purposes of studying the role of inventories in a broad perspective facing global phenomena of acculmulating crude oil inventories, or petroleum as a whole. In spite of this subtle difference, we use the term "refinery" interchangeably for the facile understanding of our theoretical model and its implication.

The optimality conditions for the representative producer are summarized as,[4]

$$\alpha e^{z_t} \left(e^{z_t} q_t\right)^{\alpha-1} \left\{1 - \exp\left(-\theta_t \cdot i_{t-1}\right)\right\} = S_t, \tag{1.2}$$

$$E_t \left[\Lambda_{t+1} e^{-r_{t+1}} \theta_{t+1} \exp\left(-\theta_{t+1} \cdot i_t\right) \left(e^{z_{t+1}} q_{t+1}\right)^{\alpha}\right] = S_t \left\{1 + c_1 i_t\right\} - E_t \left[\Lambda_{t+1} e^{-r_{t+1}} S_{t+1}\right]. \tag{1.3}$$

---

[3]In a standard macro model, we have $\Lambda_{\tau} = \frac{\beta \cdot U'(c_{\tau+1})}{U'(c_{\tau})} e^{r_{\tau}}$, where $U\left(\cdot\right)$ represents the agent's utility function and r_{τ} is the interest rate.

[4]Specifically, the values of $\left\{q_t^U, i_t\right\}_{t=1}^{\infty}$ are determined as functions of $\{S_t, z_t, \Lambda_t\}_{t=1}^{\infty}$ from the optimality conditions following (1.2) and (1.3). $\left\{q_t^P\right\}_{t=1}^{\infty}$ is determined by the accounting identity in (1.1) and optimal decisions following (1.2) and (1.3).

Equation (1.2) is the optimality condition associated with the refinery buying one more barrel of crude oil, and using the crude oil immediately for its current production. The resource demand is determined where the marginal product equals the marginal cost. Equation (1.3) is the optimality condition for the storage decision, in which the refinery equates expected marginal benefits with marginal costs of holding additional barrel of crude oil inventory. A positive value of the expected marginal benefits, left hand side in (1.3), introduces the convenience yield as directly increasing the refinery's future production capabilities. The representative producer has an incentive of holding positive level of inventory despite an expected capital loss in the future.

1.2.4 The Crude Oil Spot Price, Aggregation and Seasonality

In order to solve the proposed equilibrium model, it is necessary to deal with three issues: the crude oil spot price, aggregation and seasonality. This section provides explanations for approaches adopted in this paper.

We model the crude oil spot price process following the long-term/short-term model of [94] because of the model's flexibility to capture observed behaviors in the crude oil prices (e.g., momentum and mean-reversion). In this empirical application, we adapted their model to a discrete time stochastic process in order to obtain the closed-form forecasts for the future crude oil spot price.

Suppose that the logarithm of crude oil spot price $(\log S_t)$ consists of the long-term trend (ξ_t) and the short-term deviation (χ_t). The long-term trend follows the random walk process, reflecting the idea that oil prices themselves behave like a random walk at each time period. On the other hand, the short-term deviation from the long-term trend follows the mean-reverting AR(1) process. More

specifically,

$$\ln S_{t+1} \equiv \xi_{t+1} + \chi_{t+1},$$

$$\xi_{t+1} = \xi_t + \epsilon_{2,t+1},$$

$$\chi_{t+1} = k \cdot \chi_t + \epsilon_{3,t+1},$$

where $k \in (-1, 1)$ is the mean-reversion parameter. $\epsilon_{2,t+1}$ and $\epsilon_{3,t+1}$ are innovation processes with $E[\epsilon_{2,t+1}] = E[\epsilon_{3,t+1}] = 0$, $Var[\epsilon_{2,t+1}] = \sigma_2^2$, $Var[\epsilon_{3,t+1}] = \sigma_3^2$ for $\forall t$. Though not serially uncorrelated with their own processes, innovations are correlated to each other only at the same time period, i.e. $corr[\epsilon_{2,t+1}, \epsilon_{3,s+1}] = \rho$ for $\forall t = s$.

Given the crude oil spot price process, we adopt the risk-neutral valuation framework[5] and solve the representative producer's forecasting problem in the storage decision (1.3) as if he is risk-neutral. For stochastic processes of state variables (z_t, ξ_t, χ_t) associated with the risk-averse producer's optimal decisions, we specify risk-neutral processes by subtracting risk-premiums from underlying stochastic processes. Risk-premiums are equilibrium prices for risks that the producer pays for his hedging activities in the crude oil market. This form of risk adjustment is frequently adopted in the literature. (See, e.g., Schwartz and Smith 2000; Hamilton and Wu 2014). Denoting by $\Upsilon \equiv \left[\widetilde{\lambda}_z, \widetilde{\lambda}_\xi, \widetilde{\lambda}_\chi\right]'$ risk-premiums, we assume zero, constant, and a linear state-dependent risk premiums on the technology (z_t), long-term trend (ξ_t) and short-term deviation processes (χ_t) respectively, that is, $\widetilde{\lambda}_z \equiv 0$, $\widetilde{\lambda}_\xi = \lambda_\xi$, $\widetilde{\lambda}_\chi \equiv \lambda_\chi + \varpi \chi_t$ with $\lambda_\xi > 0$, $\lambda_\chi > 0$ and ϖ being constants.

---

[5]Implicit assumptions are the deterministic interest rate and the redundancy of the futures contract. The former guarantees the price equivalance between futures and forward contracts. More importantly, the latter validates the proposed approach of specifying risk-neutral processes by subtracting the risk-premiums from underlying processes. See more details for the application of the risk-neutral valuation framework in Duffie [Dynamic Asset Pricing Theory, 37, pp. 167-174]

Hence, risk-neutral stochastic processes are of the form:

$$\begin{bmatrix} z_{t+1} \\ \xi_{t+1} \\ \chi_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \\ -\lambda_\xi \\ -\lambda_\chi \end{bmatrix} + \begin{bmatrix} 1 & 1 & k^Q \end{bmatrix} \begin{bmatrix} z_t \\ \xi_t \\ \chi_t \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t+1}^Q \\ \epsilon_{2,t+1}^Q \\ \epsilon_{3,t+1}^Q \end{bmatrix}$$

where $k^Q \equiv k - \varpi \in (-1, 1)$[6]. Assuming that $\epsilon_{1,t+1}^Q$ is independent of both $\epsilon_{2,t+1}^Q$ and $\epsilon_{3,t+1}^Q$, innovations for risk-neutral stochastic processes $(\epsilon_{1,t+1}^Q, \epsilon_{2,t+1}^Q, \epsilon_{3,t+1}^Q)$ have identical properties as explained earlier. We provide detailed explanations for forecasting properties of the above risk-neutral processes, and for the closed-form crude oil spot price forecasts in Appendix 1.A.1.

The theory presented so far applies to one representative producer, however, available observations are aggregated at the global level. Given the trend growth observed from the global crude oil production, consumption and inventories, we assume that the number of representative producers increases over time at a fixed rate. Let N_t be the number of representative producers at period t and g be the exogenously determined average growth rate of the representative producer, i.e. $N_t = (1 + g) N_{t-1}$. The aggregate resource demands (Q_t) and inventories (I_t) become,

$$Q_t = N_t \cdot q_t$$
$$I_t = N_t \cdot i_t$$

where Q_t and I_t will be associated with globally observed aggregate quantities of the crude oil consumed and inventories. We also have the aggregate production function $F(\cdot) = N_t \cdot f(\cdot)$ and the aggregate cost function $C(\cdot) = N_t \cdot c(\cdot)$ where the representative producer's production $f(\cdot)$ and the cost function $c(\cdot)$ are as defined earlier[7].

---

[6]In the proposed model, ϖ cannot be identified separately from k. Hence, we provide the explanation based on k^Q rather than k henceforth.

[7]This approach implicitly assumes competitiveness in the global refinery industry following

Lastly, we introduce seasonally varying utilization parameters in order to deal with the strong seasonality observed both in the crude oil consumption and inventories. Specifically, we conjecture that there exist seasonal variations in the representative producer's benefit, accordingly seasonally varying convenience yield. In the northern hemisphere, for example, the aggregate demands for gasoline increase during the summer for traveling and the winter for heating purpose. In most cases, refiners are able to predict these seasonally varying aggregate demands and tend to accumulate inventories in advance when it is profitable to meet increasing demands for gasoline by utilizing their production facilities efficiently. In other words, it is reasonable to impose seasonal variations in the representative producer's benefits from holding inventories for the future purpose. We propose seasonal variations of the respective utilization parameters as follows,[8]

$$
\theta_t = \begin{cases}
\theta_1 & \text{if } t \text{ is March, April, May} \\
\theta_2 & \text{if } t \text{ is June, July, August} \\
\theta_3 & \text{if } t \text{ is September, October, November} \\
\theta_4 & \text{if } t \text{ is December, January, February}
\end{cases}
$$

1.2.5 Equilibrium Prediction

With approaches introduced in the previous section, we can yield the equilibrium prediction for the aggregate resource demands and inventories. Providing the detailed derivations in Appendix 1.A.2, the optimality conditions for the representative producer in (1.2) and (1.3) become,

$$
\alpha e^{z_t} \left(e^{z_t} Q_t\right)^{\alpha-1} (N_t)^{1-\alpha} \left(1 - \exp\left(-\theta_t \frac{I_{t-1}}{N_{t-1}}\right)\right) = S_t, \tag{1.4}
$$

---

[39] and [85].

[8]In Section 5.3, we consider alternative approaches for capturing apparent seasonality in crude oil inventories.

$$\alpha^{\phi_1}\theta_{t+1}\left(1+g\right)\left(1-\exp\left(-\theta_{t+1}\frac{I_t}{N_t}\right)\right)^{\phi_1}\exp\left(-\theta_{t+1}\frac{I_t}{N_t}\right) \tag{1.5}$$

$$\times\exp\left\{-r_{t+1}+\phi_1\left(z_t-\xi_t-k^Q\chi_t+\lambda_\xi+\lambda_\chi\right)+\frac{1}{2}\phi_1^2\left(s_1^Q+\sigma_1^2\right)\right\}$$

$$=\quad\exp\left(\xi_t+\chi_t\right)\left[1+c_1\frac{I_t}{N_t}-e^{-r_{t+1}}\exp\left\{-\left(1-k^Q\right)\chi_t-\lambda_\xi-\lambda_\chi+\frac{1}{2}s_1^Q\right\}\right],$$

where $\theta_t, \theta_{t+1} > 0$ are utilization parameters associated with the production in the current period and the storage decision for the next period respectively[9]. $\phi_1 \equiv \frac{\alpha}{1-\alpha} > 0$ and $s_1^Q \equiv \sigma_2^2 + \sigma_3^2 + 2\rho\sigma_2\sigma_3$ is the conditional variance for the one-period-ahead forecast of the crude oil spot price under the risk-neutral processes. In the aggregate resource demand equation (1.4), the number of the representative producers at previous period (N_{t-1}) works as "positive externality", yet the number at current period (N_t) works as "negative externality". This confirms our intuition under the limited storage facilities; an increasing competition tends to raise marginal storage costs in the short-run due to congestion, but it also motivates growths in storage.technologies in the long-run.

We propose to measure the equilibrium convenience yield, $CY\left(\xi_t, \chi_t, z_t, I_t\right)$, from the left hand side of the equation (1.5). After rearranging this, we have,

$$CY\left(\xi_t, \chi_t, z_t, I_t\right) \quad = \quad \Pi_0 \cdot \theta_{t+1}\left(1-\exp\left(-\theta_{t+1}\frac{I_t}{N_t}\right)\right)^{\phi_1} \tag{1.6}$$

$$\times\exp\left\{-r_{t+1}-\theta_{t+1}\frac{I_t}{N_t}+\phi_1\left(z_t-\xi_t-k^Q\chi_t\right)\right\},$$

where $\Pi_0 \equiv \alpha^{\phi_1}\left(1+g\right)\exp\left\{\phi_1\left(\lambda_\xi+\lambda_\chi\right)+\frac{1}{2}\phi_1^2\left(s_1^Q+\sigma_1^2\right)\right\}$ is a constant as a function of model parameters. Note that the proposed convenience yield is time-varying (Gibson and Schwartz 1990), and is inversely related to the current inventories (Pindyck 1994[10]), indicating that inventory holders have larger benefits

---

[9]For example, θ_1 are used for both θ_t and θ_{t+1} when the period t corresponds to March. However, we use θ_2 in place of θ_{t+1} when the period t corresponds to May. It is because it becomes summer in the next period while it is currently spring.

[10][84] proposed a net convenience yield measure from a producer's marginal costs in the storage market. Leaving out costs for storage facilities, the net convenience yield is calculated as the price difference between two adjacent futures contracts. Despite simple, such a net convenience

when inventories are relatively low. Moreover, the proposed convenience yield exhibits asymmetric responses to the long- and short-term components within the crude oil spot process.

Before concluding this section, we provide three sets of observation equations as linear functions of state variables (z_t, ξ_t, χ_t) as our approach of implementing the proposed theoretical model empirically. Detailed derivations are provided in Appendix 1.A.1 (prices for the crude oil futures contracts) and in Appendix 1.A.2 (the aggregate resource and inventory demands).

The first two observation equations are readily available from equilibrium decisions for the aggregate resource demands and inventory demands. Consider the aggregate resource demands from (1.4). Taking the log and subtracting the previous aggregate resource demands ($\log Q_{t-1}$) from the resulting equation yields,

$$
\begin{aligned}
\log\left(Q_t/Q_{t-1}\right) \;=\; & \frac{\alpha}{1-\alpha}z_t - \frac{1}{1-\alpha}\xi_t - \frac{1}{1-\alpha}\chi_t \\
& + \frac{1}{1-\alpha}\log\alpha\left(1 - \exp\left(-\theta_t\frac{I_{t-1}}{N_{t-1}}\right)\right) + \log N_t - \log Q_{t-1},
\end{aligned} \tag{1.7}
$$

where a change in the aggregate resource demands is represented as a linear function of the state variables (z_t, ξ_t, χ_t).

Next, consider the log of the aggregate inventory demands (1.5) and linear approximations of non-linear terms in the resulting equation. Linear approximations are taken for $\log\left(I_t\right)$ being around $\log\left(I_{t-1}\right)$, and for χ_t being close to $\frac{-\lambda_\chi}{1-k^Q}$ since χ_t is mean-reverting around $\frac{-\lambda_\chi}{1-k^Q}$. Rearranging this, we obtain a linear relationship between a change in the aggregate demands for inventories (I_t) and state variables (z_t, ξ_t, χ_t) as,

$$
\log\left(I_t/I_{t-1}\right) = \Phi_1\left(x_t\right)\cdot z_t + \Phi_2\left(x_t\right)\cdot\xi_t + \Phi_3\left(x_t\right)\cdot\chi_t + \Phi_4\left(x_t\right), \tag{1.8}
$$

---

yield is heavily influenced by the price movement of the nearest maturity futures contracts, rather than representing inventory holders' marginal benefit. Inferred net convenience yield in generally smaller in magnitude compared to the equilibrium convenience yield presented in this paper. Based on the similar approach, [40] estimates the net marginal convenience yield using crude oil futures prices.

where $x_t \equiv (r_{t+1}, I_{t-1})$ denotes a vector of observations that are either exogenous or predetermined at period t. Coefficients $(\Phi_1(x_t), \Phi_2(x_t), \Phi_3(x_t), \Phi_4(x_t))$ are functions of x_t and parameters as follows:

$$
\begin{aligned}
\Phi_0(x_t) &= \frac{I_{t-1}}{N_t} \left\{ \frac{\phi_1 \theta_{t+1} \exp\left(-\frac{\theta_{t+1}}{N_t} I_{t-1}\right)}{1 - \exp\left(-\frac{\theta_{t+1}}{N_t} I_{t-1}\right)} - \frac{c_1}{1 + c_1 \frac{I_{t-1}}{N_t} - \exp\left(-\lambda_\xi + \frac{1}{2} s_1^Q - r_{t+1}\right)} \right\} \\
&\quad - \theta_{t+1} \frac{I_{t-1}}{N_t} \\
\Phi_1(x_t) &= \Phi_0(x_t)^{-1} \times (-\phi_1) \\
\Phi_2(x_t) &= \Phi_0(x_t)^{-1} \times (1 + \phi_1) \\
\Phi_3(x_t) &= \Phi_0(x_t)^{-1} \times \left\{ 1 + \phi_1 k^Q + \frac{\left(1 - k^Q\right) \exp\left(-\lambda_\xi + \frac{1}{2} s_1^Q - r_{t+1}\right)}{1 + c_1 \frac{I_{t-1}}{N_t} - \exp\left(-\lambda_\xi + \frac{1}{2} s_1^Q - r_{t+1}\right)} \right\} \\
\Phi_4(x_t) &= \Phi_0(x_t)^{-1} \times \left[\frac{\theta_{t+1} I_{t-1}}{N_t} - \phi_1 \ln\left\{ 1 - \exp\left(-\frac{\theta_{t+1}}{N_t} I_{t-1}\right) \right\} + r_{t+1} \right. \\
&\quad + \ln\left\{ 1 + c_1 \frac{I_{t-1}}{N_t} - \exp\left(-\lambda_\xi + \frac{1}{2} s_1^Q - r_{t+1}\right) \right\} - \phi_1 \left(\lambda_\xi + \lambda_\chi\right) \\
&\quad + \frac{\lambda_\chi \exp\left(-\lambda_\xi + \frac{1}{2} s_1^Q - r_{t+1}\right)}{1 + c_1 \frac{I_{t-1}}{N_t} - \exp\left(-\lambda_\xi + \frac{1}{2} s_1^Q - r_{t+1}\right)} - \ln\left\{ \alpha^{\phi_1} \theta_{t+1} (1 + g) \right\} \\
&\quad \left. - \frac{1}{2} \phi_1^2 \left(s_1^Q + \sigma_1^2\right) \right]
\end{aligned}
$$

Lastly, we use closed-form spot price forecasts for evaluating prices of crude oil futures contracts with various maturities. To see this, consider the risk-neutral stochastic processes in section 1.2.4. Under the risk-neutral valuation paradigm, the "no arbitrage" price of the contingent claims coincides with the expected future cash flows under the risk-neutral stochastic process. Note that no cash payment is made at the time when a futures contract is traded. Hence, the price of the crude oil futures contract with 1 month to maturity $(F_{t,1})$ coincides with the 1-period

ahead spot price forecast under the risk neutral measure as,

$$
\begin{aligned}
F_{t,1} &= E_t^Q [S_{t+1}] \\
&= \exp\left\{ \xi_t + k^Q \cdot \chi_t - \lambda_\xi - \lambda_\chi + \frac{1}{2}s_1^Q \right\},
\end{aligned}
$$

where $k^Q, \lambda_\xi, \lambda_\chi, s_1^Q$ are as defined earlier. Taking log on this yields,

$$
\log F_{t,1} = \xi_t + k^Q \cdot \chi_t - \lambda_\xi - \lambda_\chi + \frac{1}{2}s_1^Q
$$

Similarily, the price of the crude oil futures contract with τ periods to maturity $(F_{\tau,t})$ is obtained recursively as,

$$
\log F_{t,\tau} = \xi_t + \left(k^Q\right)^\tau \cdot \chi_t - \tau \cdot \lambda_\xi - \lambda_\chi \cdot \frac{1 - \left(k^Q\right)^\tau}{1 - k^Q} + \frac{1}{2}s_\tau^Q,
$$

where $s_\tau^Q \equiv \tau \cdot \sigma_2^2 + \frac{1 - \left(k^Q\right)^{2\tau}}{1 - (k^Q)^2} \cdot \sigma_3^2 + 2\rho\sigma_2\sigma_3 \cdot \frac{1 - \left(k^Q\right)^\tau}{1 - k^Q}$ is the conditional variance associated with τ-periods ahead forecast of the crude oil spot price. Subtracting the log of the same crude oil futures contract price in the previous period $(\log F_{t-1,\tau+1})$, we obtain a linear relationship between a periodic return of the crude oil futures contracts with τ periods to maturity and state variables (ξ_t, χ_t) as,

$$
\log\left(F_{t,\tau}/F_{t-1,\tau+1}\right) = \xi_t + \left(k^Q\right)^\tau \cdot \chi_t - \tau \cdot \lambda_\xi - \lambda_\chi \cdot \frac{1 - \left(k^Q\right)^\tau}{1 - k^Q} + \frac{1}{2}s_\tau^Q - \log F_{t-1,\tau+1} \quad (1.9)
$$

From the representative producer's optimality conditions, we have three sets of observation equations for estimation, namely, (1.7), (1.8), and (1.9). In the following section, we estimate model parameters by using crude oil inventories, consumptions and futures contract prices for $1, \cdots, 12$ months to maturity[11].

---

[11]With $\tau = \{1, \ldots, 12\}$, we expect the set of equations following (1.9) captures the term structure in the crude oil futures prices, contributing to precise parameter estimates in the representative producer's production and cost functions.

1.3   Empirical Results

Given the set of observations that are linear functions of state variables, we can estimate associated model parameters by maximum likelihood using Kalman filter[12]. Five datasets are used for estimation; OECD total petroleum stocks (OECD inventory), world crude oil production, U.S. CPI all items, U.S. LIBOR, and light sweet crude oil (a.k.a WTI crude oil) futures contract prices. The monthly historical data for OECD inventory is provided by International Energy Agency and it contains total petroleum stock data for the member countries of the OECD. In order to obtain the monthly crude oil consumption, world crude oil production estimates from U.S. Energy Information Administration is used together with the monthly changes in OECD inventory. Historical daily prices for the light sweet crude oil futures contracts[13] are obtained through datastream. In order to avoid the potential liquidity issue,[14] prices for the first calendar date of each month are used for the construction of monthly price series. The monthly U.S. LIBOR is used for representing the effective historical riskfree interest rate. Lastly, the U.S.CPI, provided by OECD, is used for deflating nominal prices of the crude oil futures contracts into real prices as well as adjusting nominal interest rates for realized inflation. The time period of the analysis is from March 1989 to November 2014, where the beginning period of the analysis is determined from the availability of prices for 1 year crude oil futures contract and the ending period is determined from the availability of the OECD Inventory.

---

[12]Details regarding the state-space representation and the Kalman filter for maximum likelihood estimation can be found in Appendix 1.A.3. See more detail in Hamilton [Time Series Analysis, 62].

[13]They are traded in New York Mercantile Exchange (NYMEX) and physical delivery is required as it matures during the delivery month.

[14]Trading in the current delivery month shall cease on the third business day prior to the twenty-fifth calendar day of the month preceeding the delivery month.

1.3.1 Results

Table 1.1 reports maximum likelihood estimates. Estimates of the model parameters are consistent with the theoretical restrictions and historical observations. As for the estimates themselves, several points stand out. First, recall that we model the representative producer's problem of producing consumption goods. The output share (α) estimate indicates that crude oil resources account for approximately 31% of the consumption good production globally. Second, when the real price of crude oil was \$99.21 per barrel and the inventory level was 4,138.71 million barrels per day in January 2008, estimates for storage costs functions indicate that the marginal storage cost necessary for increasing an additional barrel of crude oil inventory is approximately \$0.51 per barrel holding other things being equal. Third, all seasonal utilization parameters ($\theta_1, \theta_2, \theta_3, \theta_4$) are statistically significant. The smallest difference among utilization parameters is 0.0011 between summer (θ_2) and fall (θ_3). Wald statistic for testing the null hypothesis of $H_0 : \theta_2 = \theta_3$, is 5.73 with the corresponding p-value being 0.0167. All seasonal utilization parameters are different from each other, and test results are provided for all pairs and jointly in Table 2. While the smallest difference is about 0.001, it is shown that holding other things being equal, 0.001 decrease in θ yields \$0.16 per barrel increase in the convenience yield during July 2003 when the inventory level of 3,935.06 millions barrels per day corresponds to the median level[15]. These estimates indicate that utilizations for summer and fall are higher than those for spring and winter since the smaller θ implies production technologies' higher dependence on available inventories. Fourth, the positive mean reversion (k^Q) indicates that the short-term deviation of the crude oil spot price is highly persistent under the risk-neutral measure. Given the slowly moving long-term trend and the positive correlation (ρ^Q) between two underlying processes, this further implies the highly persistent crude oil spot process under the risk-neutral measure. Fifth,

---

[15]0.001 decrease in θ yield \$0.13 per barrel increase in the convenience yield during March 1989 and \$0.30 per barrel increase during August 2010, where those months correspond to the period with the minimum (former) and the maximum (latter) inventory.

the long-term risk premium (λ_ξ) is positive and statistically significant, indicating risky long-term trend component in the crude oil price movement. Given the highly persistent oil price, producers need to provide counterparties with sufficient compensations for hedging their price risks, especially being associated with the long-term risk factor. On the other hand, the short-term risk premium (λ_χ) is negative, yet statistically insignificant. Accordingly, arbitrageurs do not necessarily require high risk premium when providing liquidities for hedgers when facing the short-term risk factor.

Figure 1.1 plots the historical decomposition of the crude oil spot price with the long-term trend process (Panel 1) and the short-term deviation process (Panel 2). The model fitness is shown in Panel 3 by overlaying the model forecasts with the real price for the nearest maturity crude oil futures contract (henceforth, crude oil price) on a log scale. In the historical decomposition, previous episodes in the crude oil market is shown through the lens of the long-term and short-term processes; the long-term trend represents the slowly moving trend while the short-term deviation is a mean-reverting process around the long-run average ($\frac{-\widehat{\lambda}_\chi}{1-\widehat{k}^Q} = 0.07$). Interestingly, the long-term trend has been increasing from the beginning of 2004 to the first half of 2008, reaching at the higher equilibrium level compared to the previous period[16]. Observations are generally consistent with previous episodes documented in the earlier literature (See Hamilton 2009; Kilian 2009; Kilian and Murphy 2014).

Panel 1 in Figure 1.2 plots the equilibrium convenience yield ratio, calculated from dividing the equilibrium convenience yield by the crude oil spot price (CY_t/S_t, %). The equilibrium convenience yield ratio fluctuates considerably along the business cycle, repeating the same patterns around the past recession and recovery periods. When the global economy began recovering from the early 2000's recession, for example, the equilibrium convenience yield ratio sharply increases from 14.51% in February 2002 to 19.41% in February 2003. On the other hand,

---

[16]In section 1.4.2, we revisit the potential effect of the structural change in the crude oil market from the subsample analysis.

it declines from 15.39% to 13.04% between December 2007 and June 2009 (Great Recession). Moreover, we find that the equilibrium convenience yield ratio is positively related to the crude oil price, yet is negatively related to the crude oil inventories. After the early 2000's recession, the crude oil price had increased by 59.67% and crude oil inventories had declined by 5.67%. In contrast, the oil price had decreased by 24.72% and inventories had increased by 5.57% during the Great Recession. Panel 2 plots the percentage deviation of the observed crude oil inventories from the equilibrium inventory level (Relative inventory[17]). Here, we confirm the negative relationship between equilibrium convenience yield ratios and relative inventories in general. For the two time periods considered above, the relative inventory has declined from -0.75% to -4.00% between February 2002 and February 2003, and it has increased from -0.33% to 2.92% during the Great Recession associated with decreasing equilibrium convenience yield ratios. We also find a lagged response from the relative inventory,[18] however, it is beyond the scope of this paper.

1.3.2 The Effect of Financial Investors in the Crude Oil Market

Traditionally, traders would react to releases of inventory data in a predictable way. Unexpectedly high inventories were taken as a view of weak demand and strong supplies, and the price would decline. Unexpectedly low inventories would predict an increase in the price. The fact that this relation seemed less reliable in the period of heavy financial participation led some to conclude that

---

[17][105] introduced the relative inventory as a measure for crude oil market tightness, by calculating a desired inventory level as a fitted crude oil inventory using monthly dummy variables and a linear time trend. [80] also used this for forecasting crude oil spot prices. In the following section, we address their forecasting models in detail.

[18]For relative inventory in Panel 2, local peak and trough around early 2002-2003 occurs at April 2002 and May 2003, that are 2- and 3-months lagged compared to those appeared in equilibrium convenience yield ratios respectively.

financial investors were changing the fundamental behavior of crude oil prices[19].

One description of the traditional relation between prices and inventories was developed by [105], who defined the normal inventory, NI_t as the fitted value from a regression of the observed OECD inventory OI_t on a constant, deterministic time trend, and monthly dummies:

$$OI_t = \beta_0 + \beta_1 \cdot t + \sum_{p=2}^{12} \beta_p \cdot D_p \tag{1.10}$$

where $(D_p, p = 2, \ldots, 12)$ are monthly dummies, i.e. $D_p = 1$ if the period t corresponds to the month p, $D_p = 0$ otherwise, and they further defined $RI_t \equiv OI_t - NI_t$ as the residual from this regression. Allowing for the possibility that positive and negative residuals had different effects as well as an effect of the year-over-year change $AI_t = OI_t - OI_{t-12}$, [105] proposed a forecasting model as follows,

$$F_{1,t} = a + \sum_{i=0}^{5} b_i \cdot RI_{t-i} + \sum_{i=0}^{5} c_i \cdot LI_{t-i} + d \cdot AI_t + e \cdot F_{1,t-1} + \varepsilon_t \tag{1.11}$$

where $F_{1,t}$ is the nearest maturity crude oil futures price at period t and ε_t is a regression error.

Figure 1.3 plots in-sample and out-of-sample forecasts following the traditional forecasting model. The period for in-sample forecasts coincides with the earlier literature from January 1992 to February 2001. Using estimated coefficients from in-sample forecasting period, we calculate out-of-sample forecasts from March 2001 to June 2004. The solid line in March 2001 denotes the beginning of out-of-sample forecasts. Figure 1.3 confirms the strong in-sample forecasting ability documented in the earlier literature. However, its strong forecasting ability disappears over the out-of-sample forecasting period.

Given the poor performance of the traditional forecasting model after 2003,

---

[19][48] provide a detailed review of the empirical literature investigating the role of speculation in driving the oil price after 2003.

[80] proposed the extended forecasting model by building upon the former model. After expressing RI_t in a percentage of total level ($RI\%_t$) and including CFTC non-commercial long position (X_t), the extended forecasting model is provided as follows:

$$F_{1,t} = \beta_0 + \beta_1 \cdot t + \beta_2 \cdot \Delta RI\%_t + \beta_3 \cdot RI\%_{t-1} + \beta_4 \cdot \Delta X_t + \beta_5 \cdot X_{t-1} + \eta_t \quad (1.12)$$

where $\Delta RI\%_t = RI\%_t - RI\%_{t-1}$, $\Delta X_t = X_t - X_{t-1}$ and η_t is a regression error. One thing to note is that [80] included the deterministic time trend (t) in the forecasting equation as opposed to the lagged price ($F_{1,t-1}$) considered in [105]. Figure 1.4 plots in-sample and out-of-sample forecasts following the extended forecasting model. The period for in-sample forecasts coincides with [80] from January 1992 to June 2004. Using estimated coefficients from in-sample forecasting period, we calculate out-of-sample forecasts from July 2004 to November 2014. The solid line in July 2004 denotes the beginning of the out-of-sample forecasts. Along with the extended model forecasts, we also calculate the basic model forecasts from the literature, where the latter model excludes X_t in the former forecasting equation (1.12). Figure 1.4 confirms the improved in-sample forecasting ability of the extended model as documented in [80]: the extended model exhibits the better in-sample forecasts than the basic model from 2001 to 2004 while both generated similar in-sample forecasts from 1992 to 2001.

However, when we extend the sample to include data since publication of Merino and Ortiz's study, even their model with non-commercial traders does quite badly.

On the other hand, the optimizing model developed in section 1.2 suggests that both equations (1.11) and (1.12) are misspecified. To see whether our model exhibits any of the instabilties of either (1.11) and (1.12), we performed an analogous out-of-sample evaluation. Specifically, we divide the total sample into two sub-samples: the period for in-sample forecasting covers from March 1989 to June

2004 and that for out-of-sample forecasting spans from July 2004 to November 2014. After calculating in-sample forecasts, we iterate following three steps recursively starting at June 2004 for calculating out-of-sample forecasts using parameter estimates from the first sample[20]. First, given available observations, we update the state vector and the variance matrix using the formula for updating a linear projection. Second, given updated vectors and matrices, we calculate one-period ahead forecasts with linear projections. Third, given one-period ahead forecasts, we calculate one-period ahead out-of-sample forecasts. Figure 1.5 displays forecasts of the nearest maturity crude oil futures contract price. The solid vertical line in July 2004 denotes the beginning of the out-of-sample forecasts. We find the stable crude oil price forecasting relationship both in-sample and out-of-sample. It is worth noting that out-of-sample forecasts are based on information that are either known in advance or predetermined at each forecasting period. In other words, current information is sufficient to produce a reliable forecast of the next period crude oil futures price given the stable forecasting relationship following the proposed storage model.

If equation (1.11) was misspecified, why did it seem to have such a good fit prior to 2001? To answer this question, suppose that inventories had been exactly as predicted by our model, that is, replace OI_t on the left-side of (1.10) with $\widehat{OI}_t$ the predicted equilibrium level of inventories from equation (1.8). We then reproduce the Ye, et al. exercise on this generated data, the results of which are shown in Figure 6. Even though equation (1.8) is the correct relation that exactly describes these generated data, and even though (1.10) and (1.11) are misspecified, they appear to do a reasonable job until 2004 after which they completely break down.

The explanation for the apparent success of the misspecified model in the earlier period can be found from the fact that the permanent component of crude oil prices, ξ_t, was relatively stable up until 2003, but then began to increase substan-

---

[20]MLE estimates for the first sample period is displayed in Table 1.4.

tially (see the top panel of Figure 1.1). During the period when ξ_t was relatively stable, simple relations like (1.10) did not do a bad job of fitting the data, even though they fundamentally mischaracterize the true relation between inventories and crude oil prices. On the other hand, with the large changes in ξ_t observed since 2004, the need for a better model of the relation between oil prices and inventories, such as that proposed in this paper, becomes quite evident.

1.4 Robustness Checks

In this section, we provide robustness checks in terms of concerns about endogeniety and the effect of the structural change in crude oil markets. First, we test necessary orthogonality conditions of forecasting errors in demand equations and the exogenous innovation affecting the crude oil supply. Using the lagged crude oil supply shock constructed following [73], we show that forecasting errors are not correlated with the past crude oil supply shock. This validates our maintained orthogonality assumptions. Second, after splitting the sample at June 2004, we estimate model parameters under two separate periods. While some parameter estimates indicate a potential structural change, our main estimation results are qualitatively robust across the two periods.

1.4.1 Simultaneity Issue

While estimating parameters along demand curves in (1.7) and (1.8), one concern is the endogenous nature of forecasting errors arising from simultaneous equation biases. However, given the consensus of the short-term inelastic supply of crude oil and historical episodes of the crude oil price determination, we conjecture that the primary driver of historical changes in the price of oil has been supply shocks rather than demand shocks. Specifically, what we observed historically is that 1) the crude oil price has been determined by changes in the short-term supply components by shifting the supply curve temporarily, 2) the relative size of shifts

has been larger for the supply compared to the demand curve. Details for the crude oil market structure can be found in [61] and [103].

In order to motivate our robustness check for the simultaneity issue, recall the equilibrium resource demand equation (1.7), where the representative producer forms his prediction on resource demands. Adding the forecast error ($u_{1,t}$) to "ex-ante" forecasts on resource demands, we rewrite the observation equation for resource demands as follows,

$$
\begin{aligned}
\log\left(Q_t/Q_{t-1}\right) &= \left[\frac{\alpha}{1-\alpha}z_t - \frac{1}{1-\alpha}\xi_t - \frac{1}{1-\alpha}\chi_t \right. \\
&\left. + \frac{1}{1-\alpha}\log\alpha\left(1 - \exp\left(-\theta_t\frac{I_{t-1}}{N_{t-1}}\right)\right) + \log N_t - \log Q_{t-1}\right] + u_{1,t},
\end{aligned}
$$

where terms in the square bracket correspond to the representative producer's forecasts on resource demands. If the representative producer forms his forecasts optimally given available information, then the forecasting error ($u_{1,t}$) should be uncorrelated with its own lagged values and with "ex-ante" forecasts on crude oil resource demands. From this perspective, the forecasting error ($u_{1,t}$) should be uncorrelated with innovations affecting optimal forecasts on resource demands. Using the historical crude oil supply shock (h_t) constructed following [73],[21] we test the simultaneity issue from the following implied orthogonality condition,

$$
E\left[\Delta u_{1,t} \cdot h_{t-2}\right] = 0, \tag{1.13}
$$

where $\Delta u_{1,t} \equiv u_{1,t} - u_{1,t-1}$ is the change in the forecasting error and h_{t-2} is the crude oil supply shock lagged by 2 periods. As noted in [73], this measure of the crude oil supply shock corresponds to an exogenous source of variation that

---

[21][73] proposed a structural VAR model of the global crude oil market and provided the evidence that oil price increases may have very different effects on the real price of oil, depending on the underlying cause of the price increase. Given the endogenous nature of the oil price, he proposed a methodology for decomposing unpredictable changes in the real price of oil into mutually orthogonal components such as crude oil supply shocks, aggregate demand shocks and oil specific demand shocks. We construct historical crude oil supply shocks using the real economic activity index provided in his website (http://www-personal.umich.edu/~lkilian/reaupdate.txt).

affects the crude oil price, but not innovations in crude oil demand. From this perspective, this provides a natural way of testing the endogeneity issue provided by the exogenous nature of the crude oil supply shock, i.e., the lagged crude oil supply shock (h_{t-2}) provides no information about future forecasting errors ($\Delta u_{1,s}$ for $s \geq t$) beyond that contained in current optimal forecasts on resource demands.

Similarly, we test the simultaneity issue in the inventory demand equation from the orthogonality condition implied by the exogenous crude oil supply shock (h_{t-2}). Denoting by $u_{2,t}$ the forecasting error of the representative producer's "ex-ante" inventory forecasts, we rewrite the observation equation for inventory demands by adding the forecast error ($u_{2,t}$) to "ex-ante" forecasts on inventory demands as follows,

$$\log\left(I_t/I_{t-1}\right) = \left[\Phi_1\left(x_t\right) \cdot z_t + \Phi_2\left(x_t\right) \cdot \xi_t + \Phi_3\left(x_t\right) \cdot \chi_t + \Phi_4\left(x_t\right)\right] + u_{2,t},$$

where $x_t \equiv (r_{t+1}, I_{t-1})$ denotes a vector of observations that are either exogenous or predetermined at period t and coefficients $(\Phi_1\left(x_t\right),\ \Phi_2\left(x_t\right),\ \Phi_3\left(x_t\right),\ \Phi_4\left(x_t\right))$ are same as before. The simultaneity issue in the inventory demand equation is tested from,

$$E\left[\Delta u_{2,t} \cdot h_{t-2}\right] = 0, \tag{1.14}$$

where $\Delta u_{2,t} = u_{2,t} - u_{2,t-1}$ is the change in the inventory forecasting error.

Provided by two orthogonality conditions derived as in (1.13) and (1.14), we provide the robustness check of simultaneity issues as follows. First, using estimated parameters and observations, we calculate $\{\Delta\widehat{u}_{1,t}, \Delta\widehat{u}_{2,t}\}_{t=2}^{309}$ by replacing unobservables with corresponding forecasted values of $\left\{\widehat{z}_t, \widehat{\xi}_t, \widehat{\chi}_t\right\}_{t=2}^{309}$. Simple correlation coefficient estimates with the lagged crude oil supply shock (h_{t-2}) are -0.0932 ($\Delta\widehat{u}_{1,t}$) and -0.1233 ($\Delta\widehat{u}_{2,t}$) respectively. Second, we regress $\Delta\widehat{u}_{1,t}$ (and $\Delta\widehat{u}_{2,t}$) on a constant and the lagged crude oil supply shock (h_{t-2}). In our regression specification, we also include lagged dependent variables ($\Delta\widehat{u}_{1,t-1}$ and $\Delta\widehat{u}_{1,t-1}$ respectively) and the lagged explanatory variable (h_{t-3}) for correcting effects from

serial correlations.

Table 1.3 reports estimated coefficients with p-values in paranthesis as well as the joint hypothesis test results for testing the null hypothesis of all coeffcient estimators being zero. Here, we find that both the crude oil supply shock (h_{t-2}) and its lag (h_{t-3}) do not provide statistically significant explanatory power for forecasting errors in the resource and inventory demand equations. Although we reject the null hypotheses for the joint hypothesis tests, we find that these arise from the explanatory power provided by the lagged forecasting errors in both the resource and inventory demand equations. This seems to indicate unexplained variations of forecasting errors in the demand equations, yet, these are not correlated with the exogenous variables that only affect crude oil prices. To sum up, we find that forecasting errors are not correlated with the exogenous oil supply shock, which validates our maintained orthogonality assumption.

1.4.2 Structural Change

Given the increased financial investors' participation after 2004, it has been a controversial issue whether these new participants has changed the crude oil market structure. Facing the potential structural change, we provide robustness checks by estimating structural model parameters under two sub-sample periods. The first sub-period spans from March 1989 to June 2004 (184 observations) and the second sub-period spans from July 2004 to November 2014 (125 observations). Table 1.4 reports estimates and asymptotic standard errors for two sub-periods. In general, estimates are similar to those appeared in Table 1.1. It is interesting that we find the increased autocorrelation (k^Q), correlation (ρ^Q), long-term risk-premium (λ_ξ) and the decreased short-term risk premium (λ_χ) since 2004. This seemingly reflects fundamental changes in the crude oil market as is frequently documented in the earlier literature (See, e.g., Tang and Xiong 2012; Büyükşahin and Robe 2014; Hamilton and Wu 2014). However, the stable price-inventory relationship and the cointegrated convenience yield implied from the storage model

indicate that effects from those fundamental changes are insignificant overall.

1.4.3 Inventory-Price Relation and the Recent Oil Price Plunge

After the peak of the Great Recession, the real price of crude oil revisited $90 per barrel during February 2011. Since then the volatility had reduced gradually with prices oscillating between $76.32 and $110.01 per barrel until November 2014. When the Organization of Petroleum Exporting Countries (OPEC) decided not to cut petroleum productions in the near future, the oil price started to plunge, falling by 38% in a very short period of time until March 2015. In this section, we explore the recent episode in the crude oil market by examining publicly available production and consumption forecasts based on the stable crude oil inventory-price relationship.

To begin with, consider global petroleum production and consumption forecasts measured in million barrels per day (mbpd) between August 2014 and July 2015, provided by U.S. Energy Information Administration (EIA) Short-Term Energy Outlook. Table 1.5 displays forecasts for global petroleum production (column 2) and consumption (column 3), reported at July 2014. Column 4 reports expected changes in global petroleum inventories following the accounting identity introduced in section 2.1. When the real crude oil price was $96.47 per barrel on July 2014, the global crude oil market were expected to balance over 1-year time horizon: inventories were expected to decrease at a rate of 0.52 mbpd. Next two columns report changes in production (column 5) and consumption forecasts (column 6) from each month's forecasting revision relative to those reported on July 2014. For example, the revised production forecast on August 2014 was 92.51 mbpd, reflecting a slower rate of production by 0.07 mbpd than what was forecasted on July 2014. Similarly, the revised consumption forecast on September 2014 was 92.85 mbpd, faster by 0.06 mbpd than the July 2014 forecast. The last

column reports cumulative changes in global petroleum inventories after reflecting monthly revisions in production and consumption. Note that global petroleum production forecast had been revised upward from October 2014 to April 2015, and consumption forecast for early 2015 had been revised downward. If global economies were behaving as expected on July 2014, inventories were expected to decrease at a mild rate[22] of 0.67 mbpd between January 2015 and April 2015. After reflecting monthly revisions on production and consumption, however, inventories were expected to increase rapidly at a rate of 3.79 mbpd during the first 4 months in 2015. This corresponds to 113.7 million barrels increase in inventories cumulatively [23].

In order to see whether recently revised global oil market forecasts are consistent with the recent oil price plunge, we perform an out-of-sample evaluation following the same procedure described in section 1.3.2. We use parameter estimates from the second sample in Table 1.4, and generate out-of-sample forecasts of the nearest maturity crude oil futures contract price between December 2014 and March 2015. Since world crude oil production and OECD inventory are not available for the out-of-sample evaluation period,[24] we extrapolate estimates in November 2014 by using forecasts reported in Table 1.5. Specifically, for crude oil production and consumption estimates, we apply the corresponding expected growth rate in global petroleum production and consumption during this period. Given these, monthly inventories are estimated by using the accounting identity.

Figure 1.7 displays the observed OECD inventories (solid line), their in-sample and out-of-sample forecasts (dotted line) based on U.S. EIA Short-Term

---

[22]This corresponds to a monthly average rate of 5.02 million barrels. For 309 monthly OECD petroleum inventories in our data, there are 286 observations whose monthly inventories changes (in absolute term) are greater than 5.02 million barrels. Furthermore, the average absolute monthly inventory changes is 29.57 million barrels.

[23]Over the longer time periods during previous recessions, the OECD petroleum inventories had increased by 130.52 million barrels during the early 2000 recession from March 2001 and November 2011, and by 227.52 million barrels during the Great Recession from December 2007 to June 2009. The real crude oil price had decreased by 26.67% (former), and by 24.72% (latter).

[24]The U.S. EIA reports monthly international petroleum statistics, that are generally lagged by 6 months. As of April 2014, the latest available observation is from December 2014 for crude oil production, and November 2014 for OECD inventory.

Energy Outlook. While out-of-sample forecasting period begins at December 2014, the triangle node represents inventory forecasts constructed from recursively revised global oil production and consumption forecasts. Here, we confirm that inventories are expected to increase rapidly after December 2014 when using recursively revised forecasts. For comparison, the circle node represents inventory forecasts following the July 2014 report, increasing at a slower rate than the former. Figure 1.8 displays associated crude-oil price forecasts. Though lagged, we see that the recent oil price plunge can be explained by the stable price-inventory relationship provided by our model. When only using the information available until July 2014,[25] we find that the real crude oil price is expected to be about $75 per barrel over the same horizon, that is consisistent with the real price of crude oil futures contracts maturing 1-, 2-, 3- and 4-months afterwards on November 2014.

---

[25]While using production and consumption forecasts from July 2014 report for estimating world crude oil production and OECD inventory, we replace each month's observed crude oil price by one-period-ahead out-of-sample forecast made in the previous month.

1.5   Conclusion

A recurring question in empirical work is the potential contribution of financial investors to the recent volatility of crude oil prices. This is a momentous issue for policy implications. However, empirical evidence among researchers is inconsistent and yet to indicate causal relationships. While most answers from the earlier empirical work rely on incomplete measures for representing activities of financial investors and Granger-causality tests for the analysis, we build an equilibrium storage model of the global crude oil market with crude oil inventories being treated as a separate factor of the gasoline production.

Using a newly developed structural model of the global crude oil market, we explore the role of crude oil inventories and the convenience yield arising from holding inventories. We model crude oil inventories as directly increasing the refinery's production capabilities, and show time-varying convenience yields as is consistent with the theory of storage. Our model indicates the stable forecasting relationship of prices and inventories in contrast to the typical argument of a breakdown of the forecasting relationship following increased participation by financial investors after 2004. Given this new evidence, we conclude that the potential contribution from financial investors is weaker than the magnitude public acknowledged in the crude oil market.

While the particiapation of financial investors has provided traditional players in the crude market with necessary liquidities for hedging against uncertainties, it also has been a controversial issue whether these new participants has affected significantly the crude oil market. [23] documented the increased co-movement between crude oil prices, financial assets and other commodity prices, yet this finding is inconsistent with the common knowledge about the commodity market (See Erb and Harvey 2006; Gorton and Rouwenhorst 2006; Hamilton 2009). Whether or not this increased co-movement results in the spillover effects claimed in [100], the most necessary work seems to be the welfare analysis by addressing whether or

not the participation of financial investors involves enhancing the welfare of the economy when evaluating the potential contribution of their activities in the crude oil market. Lastly, it would be interesting to investigate the effect of price uncertainties on the crude oil inventory holder's benefit. Given the existence of the role of crude oil inventories, the convenience yield is expected to be a mechanism to capure uncertainties in fundamentals as well as a mechanism to deliever these into the economy.

1.6 Acknowledgements

Chapter 1, in full, has been submitted for publication of material.

Figures and Tables
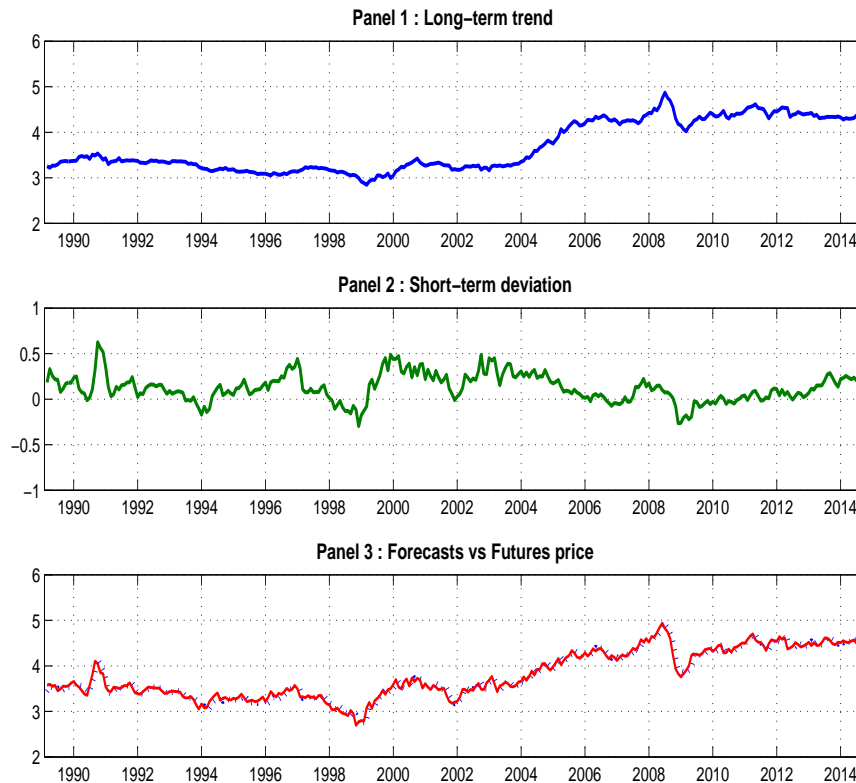
Figure 1.1: Historical Decomposition of the Crude Oil Spot Price

Figure 1.1 plots the historical decomposition of the crude oil spot price with the long-term trend process (Panel 1) and the short-term deviation process (Panel 2). Panel 3 plots the crude oil spot price (dotted line) together with the price for nearest maturity crude oil futures contract (solid line) on a log scale.
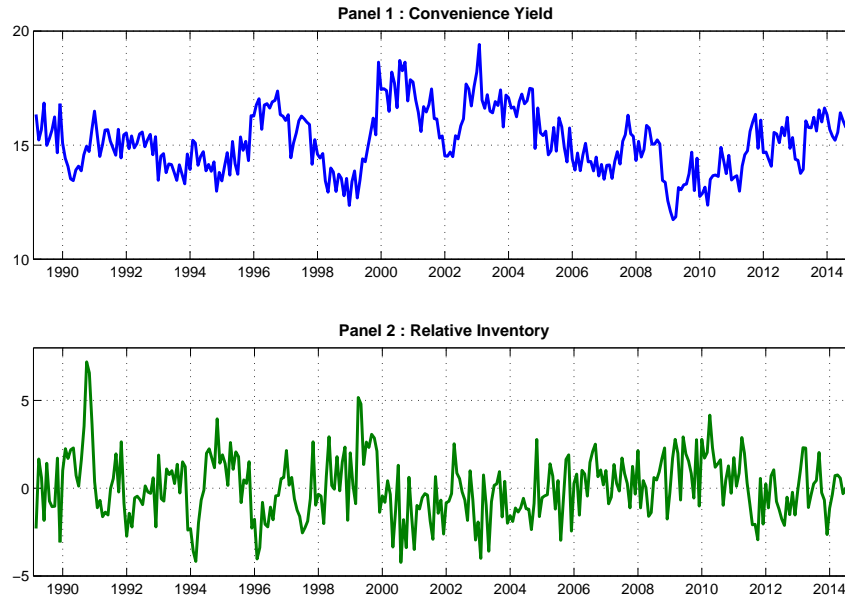
Figure 1.2: Equilibrium Convenience Yield

Figure 1.2 plots the equilibrium convenience yield in proportion to the crude oil spot price (Panel 1) and the relative crude oil inventory (Panel 2). Relative inventory is calculated as the percentage deviation of observed crude oil inventory from the equilibrium inventory inferred from the model.
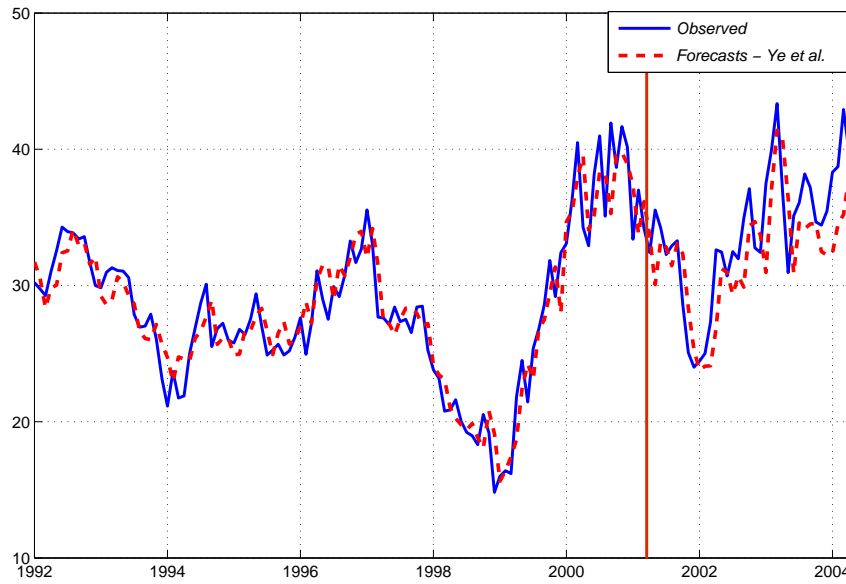
Figure 1.3: Forecasts following Ye et al. (2001)

Figure 1.3 plots in-sample and out-of-sample forecasts of the real crude oil futures price following Ye et al. (2001). While in-sample forecasts are calculated based on the period from January 1992 to February 2001, out-of-sample forecasts from March 2001 to June 2004 are calculated by using estimated coefficients from in-sample forecasting period. The solid line denotes the beginning of out-of-sample forecasts in March 2001. For comparison, observed crude oil futures prices are plotted by using the solid line.
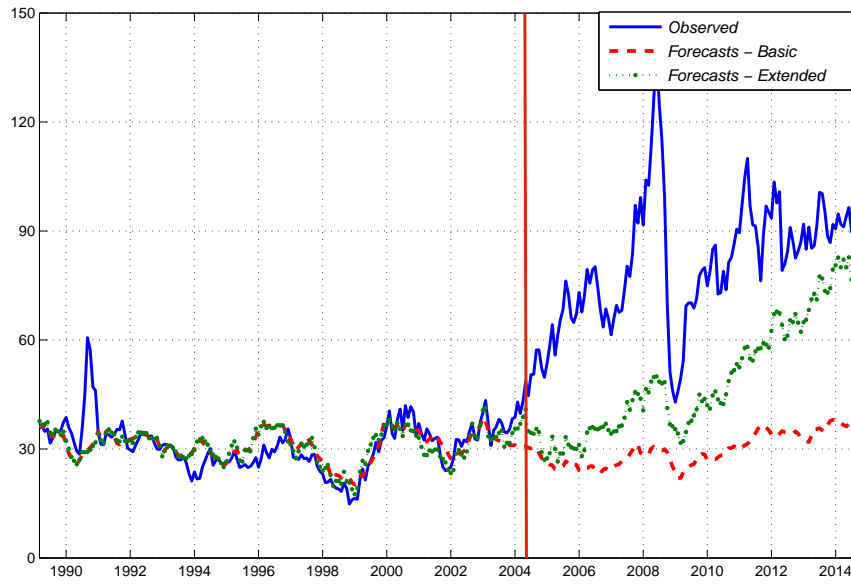
Figure 1.4: Forecasts of the Merino and Ortiz (2005) Model

Figure 1.4 plots in-sample and out-of-sample forecasts of the real crude oil futures price following Merino and Ortiz (2005). The extended (long dotted line) and the basic (short dotted line) models refer to the forecasting models with and without a role for non-commercial traders. The solid line in July 2004 denotes the beginning of the out-of-sample forecasts.
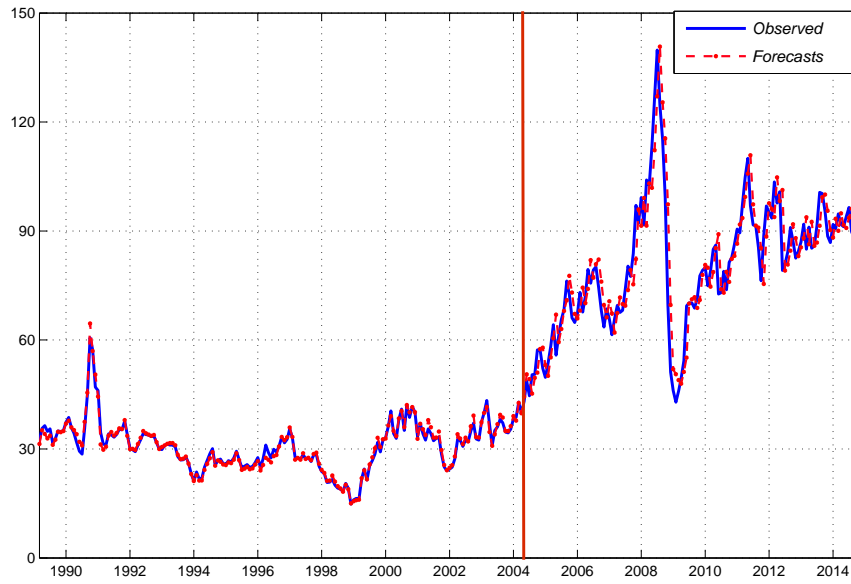
Figure 1.5: Forecasts from the Proposed Storage Model

Figure 1.5 plots in-sample and out-of-sample forecasts of the real crude oil futures price from the proposed storage model. In-sample forecasts are calculated from the proposed storage model by using observations from March 1989 to June 2004, and out-of-sample forecasts are calculated recursively from forecasts for state variables by using parameter estimates from the in-sample estimation period.
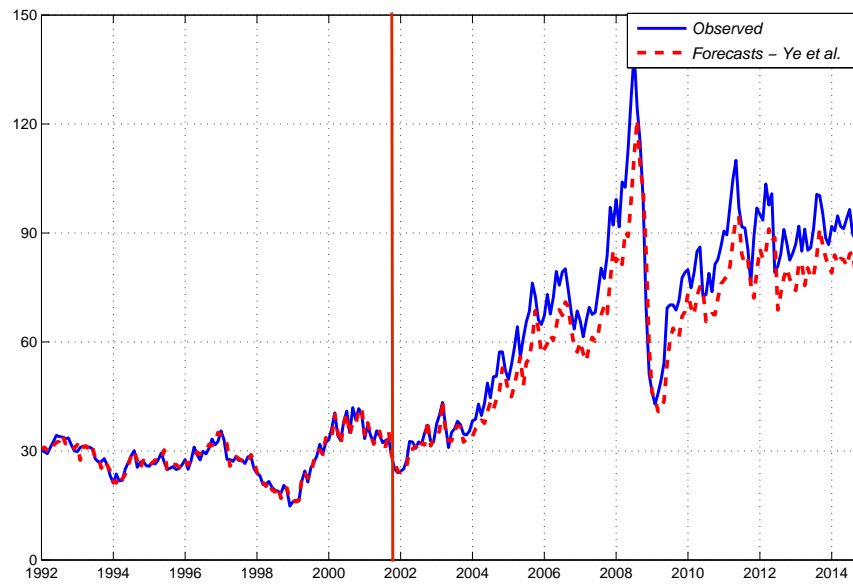
Figure 1.6: Forecasts following Ye et al. (2001) with Equilibrium Model Forecasts

Figure 1.6 plots forecasts following Ye et al. (2001) with observed crude oil inventories being replaced by the equilibrium model forecasts.
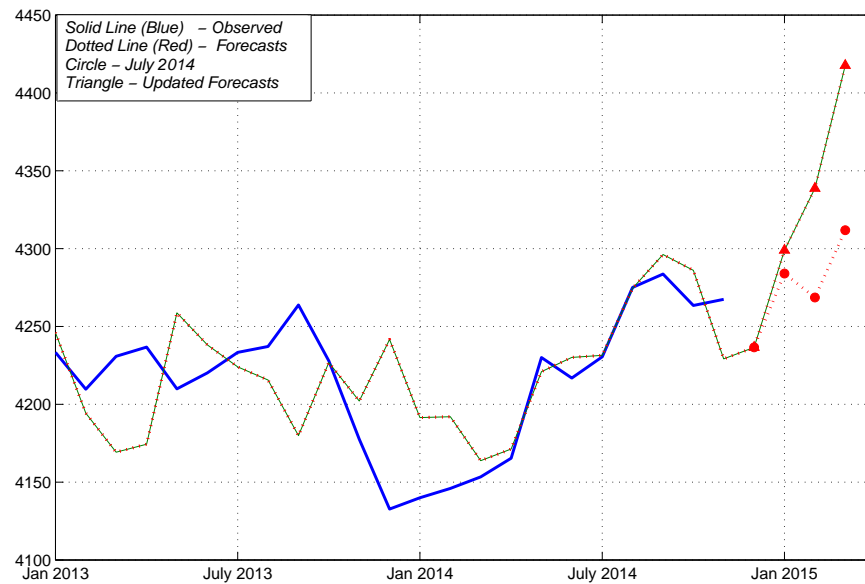
Figure 1.7: Inventory Forecasts based on EIA Short-Term Energy Outlook

Figure 1.7 plots forecasts for crude-oil inventories based on U.S. Energy Information Administration Short-term energy outlook. The solid blue line represents observed crude oil inventories and the dotted red line represents forecasts under two scenarios. While out-of-sample forecasting begins at December 2014, red circle represents out-of-sample forecasts based on July 2014 forecasts. In contrast, red triangle represents out-of-sample forecasts where monthly forecasts for petroleum production and consumption are revised over time. The vertical axis represents month-end OECD petroleum inventories in million barrel.
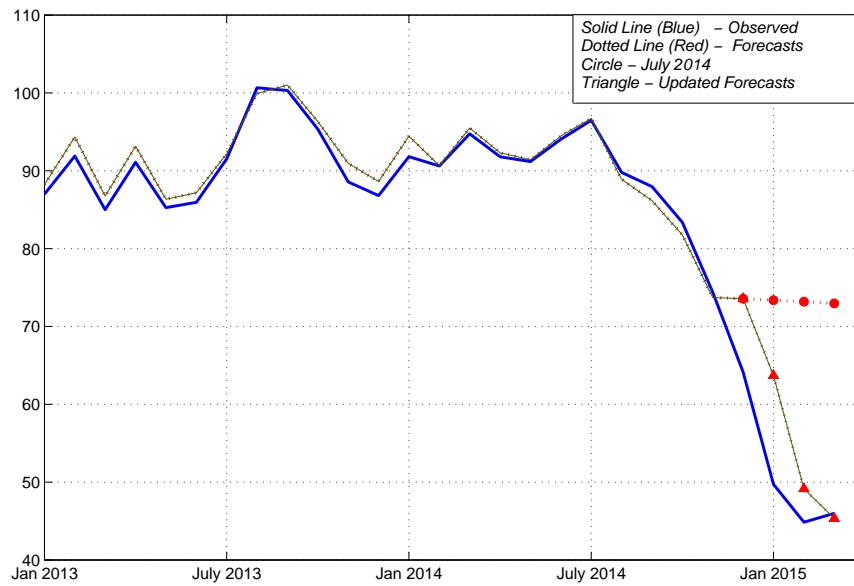
Figure 1.8: Price Forecasts based on EIA Short-Term Energy Outlook

Figure 1.8 plots forecasts for crude-oil futures contract prices based on U.S. Energy Information Administration Short-term energy outlook. The solid blue line represents observed crude oil prices and the dotted red line represents price forecasts under two scenarios. While out-of-sample forecasting begins at December 2014, red circle represents out-of-sample forecasts based on July 2014 forecasts. In contrast, red triangle represents out-of-sample forecasts where monthly forecasts for petroleum production and consumption are revised over time.

Table 1.1: MLE Estimates

Parameters	MLE	$(s.e)$	Parameters	MLE	$(s.e)$
α	0.3078	(0.0089)	σ_I	0.0176	(0.0007)
θ_1	0.0935	(0.0026)	σ_Q	0.0000	(0.0060)
θ_2	0.0896	(0.0025)	$\sigma_{F,1}$	0.0278	(0.0011)
θ_3	0.0885	(0.0025)	$\sigma_{F,2}$	0.0148	(0.0006)
θ_4	0.0917	(0.0025)	$\sigma_{F,3}$	0.0076	(0.0004)
c_1	0.0051	(0.0003)	$\sigma_{F,4}$	0.0035	(0.0002)
g	0.0003	(0.0000)	$\sigma_{F,5}$	0.0017	(0.0001)
k^Q	0.9061	(0.0019)	$\sigma_{F,6}$	0.0023	(0.0001)
ρ^Q	0.0969	(0.0601)	$\sigma_{F,7}$	0.0023	(0.0001)
λ_ξ	0.0024	(0.0002)	$\sigma_{F,8}$	0.0023	(0.0001)
λ_χ	-0.0067	(0.0664)	$\sigma_{F,9}$	0.0021	(0.0001)
σ_1	0.3095	(0.0154)	$\sigma_{F,10}$	0.0010	(0.0001)
σ_2	0.0512	(0.0021)	$\sigma_{F,11}$	0.0024	(0.0001)
σ_3	0.0777	(0.0033)	$\sigma_{F,12}$	0.0044	(0.0002)
Likelihood	14, 403.08				

Note: The likelihood function is reparametrized while preserving restrictions on parameters. The reported MLE are re-calculated from the estimated MLE and asymptotic standard errors are calculated based on initial parametric specifications. The last row reports the maximum value achieved for the log of the likelihood function.

Table 1.2: Hypothesis Tests

	H_0	W	$p-value$
(1)	$\theta_1 = \theta_2$	72.31	0.0000
(2)	$\theta_1 = \theta_3$	115.84	0.0000
(3)	$\theta_1 = \theta_4$	14.33	0.0002
(4)	$\theta_2 = \theta_3$	5.73	0.0167
(5)	$\theta_2 = \theta_4$	22.11	0.0000
(6)	$\theta_3 = \theta_4$	48.59	0.0000
(7)	$\theta_1 = \theta_2 = \theta_3 = \theta_4$	46.35	0.0000

Note: Table 1.2 reports wald test statistics and corresponding p-values for testing the null hypothesis of seasonally varying convenience yields. Specifications from (1) through (5) reports results for pairwise tests while (6) reports the result for joint test.

Table 1.3: Robustness Check - Simultaneity

(1) Resource			(2) Inventory		
Regressor	Coeff.	p-value	Regressor	Coeff.	p-value
Constant	-0.0002	0.9359	Constant	-0.0001	0.9262
h_{t-2}	-0.3398	0.7034	h_{t-2}	0.3355	0.2715
h_{t-3}	1.2692	0.1550	h_{t-3}	-0.3347	0.2750
$\nu_{1,t-1}$	-0.0401	0.0036	$\nu_{2,t-1}$	-0.0477	0.0166
F	2.5824	0.0373	F	2.3543	0.0539

Note: Table 1.3 reports regression results for testing implied orthogonal conditions. Given the exogenous crude oil supply shock, the first three columns correspond to the orthogonal condition of forecasting errors in the aggregate resource demands while the following three columns corresponds to those in the aggregate inventory demands. We include both lagged explanatory and dependent variables for taking into account potential effects from serial correlations. Lastly, we provide the joint hypothesis test result for the null hypothesis of all coefficient estimators are zero with corresponding p-values.

Table 1.4: MLE Estimates for Sub-Sample Periods

	Period 1		Period 2	
Parameters	MLE	$(s.e)$	MLE	$(s.e)$
α	0.3289	(0.0106)	0.1104	(0.0012)
θ_1	0.0843	(0.0035)	0.1295	(0.0038)
θ_2	0.0798	(0.0035)	0.1263	(0.0036)
θ_3	0.0785	(0.0034)	0.1257	(0.0035)
θ_4	0.0820	(0.0034)	0.1286	(0.0037)
c_1	0.0061	(0.0005)	0.0034	(0.0004)
g	0.0002	(0.0000)	0.0003	(0.0000)
k^Q	0.9072	(0.0028)	0.9310	(0.0021)
ρ^Q	-0.0126	(0.0843)	0.1246	(0.0966)
λ_ξ	0.0001	(0.0004)	0.0063	(0.0004)
λ_χ	-0.0031	(0.0657)	-0.0471	(0.0488)
σ_1	0.2960	(0.0183)	0.8223	(0.0530)
σ_2	0.0377	(0.0020)	0.0652	(0.0042)
σ_3	0.0909	(0.0050)	0.0622	(0.0045)
Likelihood	$8,255.93$		$6,921.95$	

Note: Table 1.4 reports MLE estimates for the baseline model under two sub-sample periods.The first sub-period (Period 1) spans from March 1989 to June 2004 (184 observations) and the second sub-period (Period 2) spans from July 2004 to November 2014 (125 observations). Although not reported, coefficient estimates for measurement errors are quantitatively similar to those in Table 1.1. The last row reports the maximum value achieved for the log of the likelihood function.

Table 1.5: Global Production and Consumption Forecasts

Dates	Forecasts (July 2014)			Revision		
	Prod.	Cons.	Inventory	Prod.	Cons.	Inventory
Aug 2014	92.58	92.18	(+0.40)	(−0.07)	(+0.00)	(+0.33)
Sep 2014	92.60	92.78	(−0.19)	(−0.23)	(+0.06)	(−0.48)
Oct 2014	92.10	92.28	(−0.18)	(+0.29)	(+0.11)	(+0.00)
Nov 2014	92.45	92.40	(+0.05)	(+0.11)	(+0.03)	(+0.13)
Dec 2014	92.11	92.24	(−0.13)	(+0.15)	(+0.03)	(−0.01)
Jan 2015	91.84	91.69	(+0.15)	(+0.16)	(−0.57)	(+0.87)
Feb 2015	92.09	92.78	(−0.68)	(+0.76)	(−0.03)	(+0.11)
Mar 2015	92.19	92.27	(−0.07)	(+1.12)	(−0.15)	(+1.20)
Apr 2015	92.51	92.58	(−0.07)	(+1.55)	(−0.13)	(+1.61)
May 2015	92.82	92.03	(+0.79)	*N/A*	*N/A*	*N/A*
June 2015	92.86	93.32	(−0.46)	*N/A*	*N/A*	*N/A*
July 2015	93.35	93.46	(−0.11)	*N/A*	*N/A*	*N/A*
Cumulative			(−0.52)			(+3.76)

Note: Table 1.5 displays global petroleum production and consumption forecasts from U.S. Energy Information Administration Short-term energy outlook, and reports the expected changes in petroleum inventories under the baseline forecasts in July 2014, and under monthly revisions. Columns 2 and 3 report global petroleum production and consumption forecasts (million barrel per day) from August 2014 to July 2015, provided in July 2014. Column 4 reports the expected average daily change in petroleum inventories as a difference between production and consumption. Column 5 and 6 report effects from monthly revisions of petroleum production and consumption relative to baseline forecasts in July 2014. The last row computes the cumulative average change in petroleum inventories, million barrel per day, under the baseline forecasts in July 2014 (column 4), and under the revised forecasts (column 7).

Appendix 1.A: Detailed Explanations

In this section, we provide detailed explanations for approaches used in the literature: the forecasting of the crude oil spot price under the risk-neutral measure, the derivation of the equilibrium prediction, and the estimation procedure with state-space representation and Kalman filter.

1.A.1 Forecasting of the Crude Oil Spot Price

Following [94], we define the logarithm of the crude oil spot price $(\log S_t)$ as the sum of the long-term trend (ξ_t) and short-term deviation processes (χ_t). Assuming risk-neutral innovation processes $(\epsilon_{2,t+1}^Q, \epsilon_{3,t+1}^Q)$ being normally distributed, the conditional expectation and the variance of $\log S_{t+1}$ under the risk-neutral processes are given by,

$$
\begin{aligned}
E_t^Q \left[\log S_{t+1} \right] &= \xi_t + k^Q \chi_t - \lambda_\xi - \lambda_\chi, \\
Var_t^Q \left[\log S_{t+1} \right] &= \sigma_2^2 + \sigma_3^2 + 2\rho\sigma_2\sigma_3.
\end{aligned}
$$

where the superscript Q indicates the calculation under the risk-neutral processes as opposed to those under the physical measure such as $E_t [\cdot]$ and $Var_t [\cdot]$.

Since the future crude oil spot price is conditionally log-normally distributed, the closed-form forecasting for the one-period-ahead crude oil spot price is provided as,

$$
\begin{aligned}
E_t^Q [S_{t+1}] &= \exp \left\{ E_t^Q \left[\log S_{t+1} \right] + \frac{1}{2} Var_t^Q \left[\log S_{t+1} \right] \right\} \\
&= \exp \left\{ \xi_t + k^Q \chi_t - \lambda_\xi - \lambda_\chi + \frac{1}{2} s_1^Q \right\}
\end{aligned}
$$

where $s_1^Q \equiv Var_t^Q \left[\log S_{t+1} \right] = \sigma_2^2 + \sigma_3^2 + 2\rho\sigma_2\sigma_3$. Similarly, it is straightforward to derive $E_t^Q \left[(S_{t+1})^{-\phi} \right] = \exp \left\{ -\phi \cdot \left(\xi_t + k^Q \cdot \chi_t - \lambda_\xi - \delta \right) + \frac{1}{2}\phi^2 s_1^Q \right\}$.

Next, the forecasting for the τ-periods ahead crude oil spot price is inferred

recursively by,

$$E_t^Q [S_{t+\tau}] = \exp \left\{ E_t^Q [\log S_{t+\tau}] + \frac{1}{2} Var_t^Q [\log S_{t+\tau}] \right\}$$

$$= \exp \left\{ \xi_t + (k^Q)^\tau \cdot \chi_t - \tau \cdot \lambda_\xi - \lambda_\chi \left(1 + k^Q + \cdots + (k^Q)^{\tau-1} \right) + \frac{1}{2} s_\tau^Q \right\}$$

where $s_\tau^Q \equiv \tau \cdot \sigma_2^2 + \frac{1-(k^Q)^{2\tau}}{1-(k^Q)^2} \cdot \sigma_3^2 + 2\rho\sigma_2\sigma_3 \cdot \frac{1-(k^Q)^\tau}{1-k^Q}$ is the conditional variance associated with τ-periods-ahead forecast of the crude oil spot price.

1.A.2 Equilibrium Prediction and Aggregate Inventory Demands

The equilibrium prediction for the aggregate inventory demands is obtained by combining the representative producer's optimal decisions with his forecasts for future crude oil spot price. Recall the optimality conditions from the model,

$$e^{z_t} Q_t = \left(\alpha e^{z_t} \left(1 - \exp \left(-\theta_t \frac{I_{t-1}}{N_{t-1}} \right) \right) \right)^{\frac{1}{1-\alpha}} (S_t)^{-\frac{1}{1-\alpha}} \cdot N_t,$$

$$e^{-r_{t+1}} E_t^Q \left[(e^{z_{t+1}} Q_{t+1})^\alpha (N_{t+1})^{1-\alpha} \frac{\theta_{t+1}}{N_t} \exp \left(-\theta_{t+1} \frac{I_t}{N_t} \right) \right]$$

$$= S_t \left(1 + c_1 \frac{I_t}{N_t} \right) - e^{-r_{t+1}} E_t^Q [S_{t+1}],$$

where we rearrange the equation for aggregate resource demands in (1.4).

After moving time subscript one-period forward for the rearranged aggregate resource demands, we plug the resulting equation into aggregate inventory demands whose left hand side becomes,

$$L(I_t, N_t, r_{t+1}) \cdot E_t^Q \left[(e^{z_{t+1}})^{\frac{\alpha}{1-\alpha}} (S_{t+1})^{-\frac{\alpha}{1-\alpha}} \right] = S_t \left(1 + c_1 \frac{I_t}{N_t} \right) - e^{-r_{t+1}} E_t^Q [S_{t+1}]$$

where $L(I_t, N_t, r_{t+1}) \equiv e^{-r_{t+1}} \alpha^{\frac{\alpha}{1-\alpha}} \theta_{t+1} (1+g) \left(1 - \exp \left(-\theta_{t+1} \frac{I_t}{N_t} \right) \right)^{\frac{\alpha}{1-\alpha}} \exp \left(-\theta_{t+1} \frac{I_t}{N_t} \right)$ does not depend on the representative producer's forecasts on the future technology process z_{t+1}, and the future spot price S_{t+1}. Assuming the independence of the technology process and the log-normal property such as $E_t^Q [(e^{z_{t+1}})^\varphi] = \exp \left(\varphi z_t + \frac{1}{2} \varphi^2 \sigma_1^2 \right)$ for any

φ under the risk-neutral measure, the equilibrium storage equation is further reduced to,

$$L\left(I_t, N_t, r_{t+1}\right) \cdot \exp\left\{-\phi_1\left(-z_t + \xi_t + k^Q \chi_t - \lambda_\xi - \lambda_\chi\right) + \frac{1}{2}\phi_1^2\left(\sigma_1^2 + s_1^Q\right) - r_{t+1}\right\}$$

$$= S_t\left[1 + c_1\frac{I_t}{N_t} - \exp\left\{-\left(1 - k^Q\right)\chi_t - \lambda_\xi - \lambda_\chi + \frac{1}{2}s_1^Q - r_{t+1}\right\}\right]$$

where $\phi_1 \equiv \frac{\alpha}{1-\alpha}$.

After taking the log of both sides, the measurement equation of the aggregate inventory demands is obtained by approximating non-linear terms in the resulting equation. Linear approximations are taken where χ_t and I_t being close to $\frac{-\lambda_\chi}{1-k^Q}$ and I_{t-1} respectively since χ_t is mean-reverting around $\frac{-\lambda_\chi}{1-k^Q}$ and I_t is close to I_{t-1} for most of sample. Denoting by $x_t \equiv (r_{t+1}, I_{t-1})$ a vector of obervations that are either exogenous or predetermined at period t, we have,

$$\log\left(I_t/I_{t-1}\right) = \Phi_1\left(x_t\right) \cdot z_t + \Phi_2\left(x_t\right) \cdot \xi_t + \Phi_3\left(x_t\right) \cdot \chi_t + \Phi_4\left(x_t\right),$$

where $\Phi_1(x_t)$, $\Phi_2(x_t)$, $\Phi_3(x_t)$ and $\Phi_4(x_t)$ are as specified in Section 3.2.

1.A.3 State-space Representation and Kalman Filter

Given the set of observations that are linear functions of state variables, we estimate associated parameters by maximum likelihood using the Kalman filter. First, we represent the system of equations in the state-space form. Let $\zeta_t = [z_t, \xi_t, \chi_t]'$ denote a (3×1) vector of unobservable state variables and y_t denote a (14×1) vector of observed variables at each period t, i.e. $y_t = (I_t, \ln Q_t, \ln F_{t,1}, \ln F_{t,2}, \ldots, \ln F_{t,12})'$. Then y_t can be described in terms of a linear function of unobservable state vector ζ_t. Denoting by x_t a vector of observations that are either exogenous or predetermined, the state-space representation is given by the following system of equations,

$$\xi_{t+1} = C + G\xi_t + v_{t+1}, \tag{1.15}$$

$$y_t = a\left(x_t\right) + H\left(x_t\right)' \xi_t + w_t. \tag{1.16}$$

where $a(x_t)$ and $H(x_t)$ in the observation equation denote (14×1) vector valued function, and (14×2) matrix valued function, those are specified below. In the transition equation, C and G denote respectively a (3×1) coefficient vector and a (3×3) coefficient matrix as,

$$
C = \begin{pmatrix} 0 \\ -\lambda_\xi \\ -\lambda_\chi \end{pmatrix} \quad
G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & k^Q \end{pmatrix}
$$

Let $\boldsymbol{y}_{t-1} \equiv \left(y'_{t-1}, y'_{t-2}, \ldots, y'_1, x'_{t-1}, \ldots, x'_1 \right)'$ denote data observed through date $t-1$. Assuming that conditional on x_t and $\boldsymbol{y}_{t-1}$, the vector $\left(v'_{t+1}, w'_t \right)'$ has the normal distribution,

$$
\begin{bmatrix} v_{t+1} \\ w_t \end{bmatrix} \mid x_t, \boldsymbol{y}_{t-1} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \right) \tag{1.17}
$$

where restrictions on the unknown matrix Q and the observation error matrix R are given as

$$
Q = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & \rho^Q \sigma_2 \sigma_3 \\ 0 & \rho^Q \sigma_2 \sigma_3 & \sigma_3^2 \end{pmatrix} \quad
R = \begin{pmatrix} \sigma_I^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_Q^2 & 0 & \cdots & 0 \\ 0 & 0 & \sigma_{F,1}^2 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_{F,12}^2 \end{pmatrix}
$$

where parameters $\sigma_1, \sigma_2, \sigma_3 > 0$ and $\rho^Q \in (-1, 1)$ are unknown parameters in state variables as described earlier. $\sigma_I^2, \sigma_Q^2 > 0$ are measurement errors for aggregate demands for inventories and resource respectively, and $\sigma_{F,\tau}^2$ is the measurement error for the price of the crude oil futures contract with the τ-periods until maturity for $\tau = 1, \ldots, 12$.

Lastly, we are left with specifying one unknown matrix $H(x_t)$ and one unknown vector $a(x_t)$, yet restrictions are obtained from three sets of observation equations de-

rived earlier. Recall these observation equations;

$$\log\left(I_t/I_{t-1}\right) \;=\; \Phi_1\left(x_t\right)\cdot z_t + \Phi_2\left(x_t\right)\cdot \xi_t + \Phi_3\left(x_t\right)\cdot \chi_t + \Phi_4\left(x_t\right)$$

$$\log\left(Q_t/Q_{t-1}\right) \;=\; \Psi_1\cdot z_t + \Psi_2\cdot \xi_t + \Psi_2\cdot \chi_t + \Psi_3\left(x_t\right)$$

$$\log\left(F_{t,\tau}/F_{t-1,\tau+1}\right) \;=\; \xi_t + \Omega_{1,\tau}\left(x_t\right)\cdot \chi_t + \Omega_{2,\tau}\left(x_t\right)$$

where $\Psi_1 \equiv \frac{\alpha}{1-\alpha}$, $\Psi_2 \equiv -\frac{1}{1-\alpha}$, $\Psi_3\left(x_t\right) \equiv \frac{1}{1-\alpha}\left\{\log\alpha\left(1-\exp\left(-\theta_t\frac{I_{t-1}}{N_{t-1}}\right)\right)\right\} + \log N_t - \log Q_{t-1}$, $\Omega_{1,\tau}\left(x_t\right) \equiv \left(k^Q\right)^\tau$ and $\Omega_{2,\tau}\left(x_t\right) \equiv -\tau\lambda_\xi - \frac{1-\left(k^Q\right)^\tau}{1-k^Q}\lambda_\chi + \frac{1}{2}s_\tau^Q - \log F_{t-1,\tau+1}$.
Lastly, $\Phi_1\left(x_t\right)$, $\Phi_2\left(x_t\right)$, $\Phi_3\left(x_t\right)$, $\Phi_4\left(x_t\right)$ are defined as earlier.

With these, we can specify $H\left(x_t\right)$ and $a\left(x_t\right)$ following

$$H\left(x_t\right)' = \begin{bmatrix} \Phi_1\left(x_t\right) & \Phi_2\left(x_t\right) & \Phi_3\left(x_t\right) \\ \Psi_1 & \Psi_2 & \Psi_2 \\ 0 & 1 & \Omega_{1,1}\left(x_t\right) \\ \vdots & \vdots & \vdots \\ 0 & 1 & \Omega_{1,12}\left(x_t\right) \end{bmatrix} \quad a\left(x_t\right) = \begin{bmatrix} \Phi_4\left(x_t\right) \\ \Psi_3\left(x_t\right) \\ \Omega_{2,1}\left(x_t\right) \\ \vdots \\ \Omega_{2,12}\left(x_t\right) \end{bmatrix}$$

which completes the state-space representation.

Next, consider the calculation of the conditional likelihood using Kalman filter for the maximum likelihood estimation. Suppose $\xi_t|\boldsymbol{y}_{t-1} \sim N\left(\widehat{\xi}_{t|t-1}, P_{t|t-1}\right)$. Since x_t contains only strictly exogenous variables or lagged values of $\boldsymbol{y}$, we also have

$$\xi_t|\boldsymbol{y}_{t-1}, x_t \sim N\left(\widehat{\xi}_{t|t-1}, P_{t|t-1}\right)$$

With (1.15),(1.16) and (1.17), it can be shown that

$$\begin{bmatrix} \xi_t \\ y_t \end{bmatrix} \mid \boldsymbol{y}_{t-1}, x_t$$

$$\sim\; N\left(\begin{bmatrix} \widehat{\xi}_{t|t-1} \\ a\left(x_t\right) + \left[H\left(x_t\right)\right]'\widehat{\xi}_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1}H\left(x_t\right) \\ H\left(x_t\right)'P_{t|t-1} & H\left(x_t\right)'P_{t|t-1}H\left(x_t\right) + R \end{bmatrix}\right)$$

where $a\left(x_t\right)$ and $H\left(x_t\right)$ are treated as deterministic conditional on x_t. Since they are

jointly normally distributed, the conditional distribution of ξ_t given $\boldsymbol{y}_{t-1}, x_t, y_t$ becomes,

$$\xi_t \mid \boldsymbol{y}_{t-1}, x_t, y_t \sim N\left(\widehat{\xi}_{t|t}, P_{t|t}\right)$$

where we used the formula for updating a linear projection;

$$\widehat{\xi}_{t|t} = \widehat{\xi}_{t|t-1} + P_{t|t-1}H\left(x_t\right)\left[H\left(x_t\right)' P_{t|t-1}H\left(x_t\right) + R\right]^{-1} \times \left[y_t - a\left(x_t\right) - [H\left(x_t\right)]' \widehat{\xi}_{t|t-1}\right]$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H\left(x_t\right)\left[H\left(x_t\right)' P_{t|t-1}H\left(x_t\right) + R\right]^{-1} H\left(x_t\right)' P_{t|t-1}$$

It then follows,

$$\xi_{t+1}|\boldsymbol{y}_t \sim N\left(\widehat{\xi}_{t+1|t}, P_{t+1|t}\right)$$

where $\widehat{\xi}_{t+1|t} = C + G\widehat{\xi}_{t|t}$ and $P_{t+1|t} = GP_{t|t}G' + Q$.

Next, we can calculate sample likelihood using Kalman filter as derived earlier. Assume that the initial state ξ_1 is distributed as $N\left(\widehat{\xi}_{1|0}, P_{1|0}\right)$. Given observations $\{y_t, x_t\}$ for $t = 1, \ldots, T$, the distribution of y_t conditional on $(y_{t-1}, \ldots, y_1, x_t, x_{t-1}, \ldots, x_1)$ is normal with mean $\left(a\left(x_t\right) + H\left(x_t\right)' \widehat{\xi}_{t|t-1}\right)$ and variance $\left(H\left(x_t\right)' P_{t|t-1}H\left(x_t\right) + R\right)$, that is, for $t = 1, 2, \ldots, T$, we have the conditional likelihood following

$$\begin{aligned}
f_{Y_t|X_t, Y_{t-1}}\left(y_t|x_t, \boldsymbol{y}_{t-1}\right) &= (2\pi)^{-n/2} \left|H\left(x_t\right)' P_{t|t-1}H\left(x_t\right) + R\right|^{-1/2} \\
&\times \exp\left[-\frac{1}{2}\left(y_t - a\left(x_t\right) - H\left(x_t\right)' \widehat{\xi}_{t|t-1}\right)'\right. \\
&\times \left.\left(H\left(x_t\right)' P_{t|t-1}H\left(x_t\right) + R\right)^{-1}\left(y_t - a\left(x_t\right) - H\left(x_t\right)' \widehat{\xi}_{t|t-1}\right)\right]
\end{aligned}$$

With this, the sample log likelihood becomes,

$$l\left(\Theta\right) = \sum_{t=1}^{T} \log f_{Y_t|X_t, \boldsymbol{Y}_{t-1}}\left(y_t|x_t, \boldsymbol{y}_{t-1}\right),$$

where $\Theta = \left(\alpha, \theta_1, \theta_2, \theta_3, \theta_4, k^Q, c_1, \rho^Q, \lambda_\xi, \lambda_\chi, \sigma_1, \sigma_2, \sigma_3, \sigma_I, \sigma_Q, \sigma_{F,1}, \ldots, \sigma_{F,12}\right)$ are parameters to be estimated numerically.

Appendix 1.B: Further results

Table 1.B.1 reports maximum likelihood estimates under three cases: using the US imported gasoline from Gulf, using the price of the gasoline futures contracts traded in the New York Mercantile Exchange, and without considering seasonally varying utilization parameter in the model. Compared to the main result in Table 1.1, most estimates are similar both in magnitudes and signs, yet attaining at lower likelihoods.

Table 1.B.1 : MLE Estimates Under the Alternative Approach

Parameters	(1) Gulf		(2) New York		(3) No seasonality	
	MLE	$(s.e.)$	MLE	$(s.e.)$	MLE	$(s.e.)$
α	0.1805	(0.0054)	0.1812	(0.0054)	0.3071	(0.0088)
θ	0.1196	(0.0032)	0.1179	(0.0032)	0.0900	(0.0028)
c_1	0.0038	(0.0004)	0.0040	(0.0004)	0.0053	(0.0004)
g	0.0004	(0.0000)	0.0004	(0.0000)	0.0003	(0.0000)
k^Q	0.9057	(0.0019)	0.9060	(0.0019)	0.9061	(0.0019)
ρ^Q	0.0969	(0.0600)	0.0949	(0.0600)	0.0964	(0.0601)
λ_ξ	0.0024	(0.0002)	0.0024	(0.0002)	0.0024	(0.0002)
λ_χ	0.0000	(0.0666)	-0.0068	(0.0665)	-0.0067	(0.0664)
σ_1	0.3031	(0.0286)	0.2564	(0.0264)	0.3098	(0.0154)
σ_2	0.0511	(0.0021)	0.0511	(0.0021)	0.0512	(0.0021)
σ_3	0.0778	(0.0033)	0.0778	(0.0033)	0.0778	(0.0033)
Likelihood	$14,219.79$		$14,275.11$		$14,343.85$	

Note: Table 1.B.1 reports MLE estimates under alternative approaches. The first two specifications consider gasoline prices for capturing the seasonality in crude oil inventories by using US gasoline price imported from Gulf and by using prices for gasoline futures contracts traded in the New York Mercantile Exchange. For comparison, the last column reports MLE estimates from the proposed storage model with time-invariant utilization parameter (θ). Although not reported, coefficient estimates for measurement errors are quantitatively similar to those in Table 1.1. The last row reports the maximum value achieved for the log of the likelihood function.

Chapter 2

The Usefulness of Cross-sectional Dispersion for Forecasting Aggregate Stock Price Volatility

Abstract. Does cross-sectional dispersion in the returns of different stocks help forecast volatility of the S&P 500 index? This paper develops a model of stock returns where dispersion in returns across different stocks is modeled jointly with aggregate volatility. Although specifications that allow for feedback from cross-sectional dispersion to aggregate volatility have a better fit in sample, they prove not to be robust for purposes of out-of-sample forecasting. Using a full cross-section of stock returns jointly, however, I find that use of cross-sectional dispersion can help improve parameter estimates of a GARCH process for aggregate volatility to generate better forecasts both in sample and out of sample. Given this evidence, I conclude that cross-sectional information helps predict market volatility indirectly rather than directly entering in the data-generating process.

2.1 Introduction

Modeling and forecasting volatility is an important task and a popular research agenda in financial markets. Volatility models play key roles in the academic literature for testing the fundamental tradeoff between risk and return of financial assets and for investigating the causes and consequences of the volatility dynamics in the economy. Volatility forecasts have many practical applications as well. For example, volatility forecasts are used for market timing decisions, portfolio selection, risk management and pricing of financial derivatives such as options and forward contracts since risks are measured by the volatility of the financial asset returns.

Given its importance, there is a growing number of models and approaches for forecasting volatility in financial assets. Since [43] and [13], Autoregressive Conditional Heteroskedasticity (ARCH) models are the most popular, formulating the volatility forecasts of a return as a function of known variables. Adopting the specific functional form and/or alternative explanatory variables in the volatility forecasts, there have been numerous extensions of ARCH models focusing on highlighted characteristics such as volatility persistence, asymmetry, long memory properties and a leptokurtic distribution of financial asset returns. While earlier researchers use variation over time in variables of interests, cross-sectional information began to be recognized as an important source for improving volatility forecasts. [26] and [29] suggested cross-sectional dispersion across individuals as an important source of individual stock volatility[1]. In a similar context, [69] tested whether cross-sectional dispersion helps forecast volatility of individual stock returns using GARCH-X models, where X refers to "cross-sectional dispersion". Although better specified, they found a trivial improvement from GARCH-X for out-of-sample volatility forecasts, concluding that GARCH-X models do not necessarily outperform than GARCH models in forecasting individual stock volatility.

---

[1]The contribution of cross-sectional dispersion is referred to as firm-specific (idiosyncratic) volatility in [26] and common heteroscedasticity in asset specific returns in [29].

In this paper, I investigate potential channels by which cross-sectional information might help predict aggregate volatility. I develop a model of individual returns that could be applied to the study of volatility in any financial assets, though the interest in this paper is in volatility of the stock market index such as the S&P 500 Composite Index (henceforce, S&P 500 Index). The approach includes a GARCH-X model as a special case in which measures of cross-sectional dispersion appear in the equation for predicting aggregate volatility, an approach previously investigated by [69]. I find that although such GARCH-X specifications can improve in-sample forecasting accuracy, cross-sectional dispersion does not appear to be useful for out-of-sample forecasting.

Next, I investigate another channel in which cross-sectional information helps predict aggregate volatility by providing accurate parameter estimates. After jointly modeling the full cross-section of individual stock returns, I estimate population parameters in the bivariate GARCH process for aggregate volatility and cross-sectional dispersion. While sharing the same GARCH process with univariate GARCH, I find improved forecasting accuracies that are statistically significant both in sample and out of sample for Mean Squared Error (MSE) and Mean Absolute Error (MAE) loss criteria. Using [54]'s conditional predictive ability tests, I show that by jointly utilizing cross-sectional information, it also provides more accurate out-of-sample volatility forecasts in times of recessions as well as during bear markets. I conclude that cross-sectional information helps predict market volatility indirectly insofar as it helps to obtain accurate parameter estimates for volatility forecasts.

This paper is organized as follows. The following section describes a model of stock returns and clarifies the relation between models of aggregate stock volatility and the volatility of individual returns. In section 2.3, I investigate two potential channels whereby cross-sectional information could improve volatility forecasts of the S&P 500 Index return: 1) cross-sectional dispersion as an additional explanatory variable as in GARCH-X, 2) cross-sectional dispersion as an aid in parameter

estimation when individual stock returns are jointly modeled. Section 2.4 provides robustness checks by using an alternative volatility proxy, alternative measures for cross-sectional dispersion including [69]'s cross-sectional market volatility. I also consider the potential effects from ordinary cash dividends of individual stocks and those from previously observed extreme stock market events when predicting the volatility of the S&P 500 index return. All robustness checks confirm results provided in section 2.3. Lastly, Section 2.5 concludes.

2.2   Model

Let $r_{i,t}$ denote the monthly return on individual stock i measured in percent. For example, $r_{i,t} = -1.5$ means that stock i fell 1.5% from month $t - 1$ to t. My interest is in characterizing aggregate market volatility as measured by some weighted average of individual returns,[2]

$$r_t = \sum_{i=1}^{N_{t-1}} w_{i,t-1} r_{i,t} \tag{2.1}$$

where $w_{i,t-1}$ is a predetermined weight of a stock $i's$ return in the evolution of the stock index return in period t. For example, $w_{i,t-1} = N_{t-1}^{-1}$ for an equal-weighted index and $w_{i,t-1} = p_{i,t-1} s_{i,t-1} / \sum_{j=1}^{N_{t-1}} p_{j,t-1} s_{j,t-1}$ for a value weighted index where $p_{i,t-1}$ is stock i's price and $s_{i,t-1}$ is its number of outstanding shares at time $t - 1$.

One approach would be to fit a univariate GARCH-X(1,1) model to the aggregate return:

$$r_t = \phi^0 + \phi^1 r_{t-1} + u_t, \tag{2.2}$$

$$u_t = \sigma_t \cdot \varepsilon_t, \tag{2.3}$$

$$\varepsilon_t \sim \text{ i.i.d. } (0,1),$$

---

[2]When a stock i is newly added in the stock index in period t, there is no contribution of a stock i's return $r_{i,t}$ to the stock index return r_t.

$$\sigma_t^2 = \varpi + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 + \pi x_{t-1}. \tag{2.4}$$

Here x_{t-1} is a measure of the cross-sectional dispersion of stocks at time $t-1$. Note that equation (2.4) includes the standard univariate GARCH(1,1) as a special case when $\pi = 0$.

One question is what measure of dispersion to use for x_{t-1} and what kind of model for individual stock returns would be consistent with a process like (2.4) for aggregate returns. Having an explicit answer to the latter question will also clarify the way in which data on individual stock returns might be helpful for estimating the parameters of equation (2.4). Consider an AR(1) forecasting model of stock i's return of the form,

$$r_{i,t} = \phi_i^0 + \phi_i^1 r_{i,t-1} + u_{i,t}. \tag{2.5}$$

Let v_t be an aggregate shock affecting all stock returns, distributed as $N(0, \sigma_t^2)$ conditional on available observations. Letting κ_t denote a separate shock governing the cross-section dispersion of stock returns,[3] the forecasting error of stock i's return $(u_{i,t})$ is modeled as

$$u_{i,t} = \delta_{i,t-1} [\lambda_i v_t + \kappa_t \eta_{i,t}], \tag{2.6}$$

where $\delta_{i,t-1}$ denotes a predetermined loading of stock i on the two shocks v_t and κ_t, λ_i captures the degree to which a stock i responds to v_t, and $\eta_{i,t}$ is stock i's idiosyncratic forecasting error presumed to follow a martingale difference sequence with unit variance

$$E\left(\eta_{i,t}^2 | \mathcal{F}_{t-1}\right) = 1,$$

where $\mathcal{F}_{t-1}$ denotes all observed variables through $t-1$.

One special case of interest comes from the idea that small stocks tend to

---

[3][29] and [71] suggested a strong commonality in asset specific volatilities, so that the average squared asset-specific return across a large number of stocks varies over time.

be more risky, for which we could specify $\delta_{i,t-1}$ as,

$$\delta_{i,t-1} = \frac{1}{N_{t-1}w_{i,t-1}} \tag{2.7}$$

where a stock with below-average weight $(w_{i,t-1} < 1/N_{t-1})$ is treated in (2.7) as being more exposed to the aggregate shock as well as having a higher weight on the dispersion shock κ_t. Support for such a specification can be found for example in the results of [95] and [5].

Since time variation in the cross-sectional dispersion in this model is driven by the value of κ_t, we could use κ_{t-1}^2 in place of x_{t-1} in (2.4) if we had direct observations available on κ_{t-1}.

Note that the aggregate forecasting error u_t is related to the errors in forecasting individual returns $u_{i,t}$ through the identity

$$
\begin{aligned}
u_t &= \sum_{i=1}^{N_{t-1}} w_{i,t-1}u_{i,t} \\
&= v_t \cdot \frac{1}{N_{t-1}} \sum_{i=1}^{N_{t-1}} \lambda_i + \kappa_t \cdot \frac{1}{N_{t-1}} \sum_{i=1}^{N_{t-1}} \eta_{i,t}
\end{aligned}
\tag{2.8}
$$

For N_t large, it is reasonable to assume that $\frac{1}{N_{t-1}} \sum_{i=1}^{N_{t-1}} \lambda_i \to 1$. For example, with a constant-sized sample of equal-weighted stocks, $w_{i,t} = 1/N$, in which case $N^{-1} \sum_{i=1}^{N} \lambda_i = 1$ would always exactly equal 1 by construction[4]. Likewise for N_{t-1} large, it would typically be the case that $\frac{1}{N_{t-1}} \sum_{i=1}^{N_{t-1}} \eta_{i,t} \to 0$, implying $\kappa_t \cdot \frac{1}{N_{t-1}} \sum_{i=1}^{N_{t-1}} \eta_{i,t} \xrightarrow{p} 0$. Thus when N_{t-1} is large the aggregate return u_t gives a

---

[4]Consider the simple case for the equal-weighted index return $(w_{i,t} = N^{-1}$ for $\forall i$ and $\forall t)$. Then, a consistent estimate of λ_i for each i can be obtained by a univariate regression of stock i on the equal-weighted index;, $u_{i,t} = \lambda_i \cdot u_t + e_{i,t}$ estimated by OLS for $t = 1, \ldots, T$. It is clear that $\widehat{\lambda}_i^{OLS} = \left(\sum_{t=1}^{T} u_t^2\right)^{-1} \left(\sum_{t=1}^{T} u_{i,t}u_t\right)$, satisfying $N^{-1} \sum_{i=1}^{N} \widehat{\lambda}_i^{OLS} = 1$.

direct observation on the common shock v_t :

$$p\lim_{N_t\to\infty} u_t = v_t \tag{2.9}$$

Note further from (2.6) that

$$\sum_{i=1}^{N_{t-1}} w_{i,t-1}\left(u_{i,t} - \delta_{i,t-1}\lambda_i v_t\right)^2 = \kappa_t^2 \sum_{i=1}^{N_{t-1}} w_{i,t-1}\left(\delta_{i,t-1}\eta_{i,t}\right)^2$$
$$\xrightarrow{p} \kappa_t^2 \cdot \overline{\delta}^2$$

provided $\sum_{i=1}^{N_{t-1}} w_{i,t-1}\delta_{i,t-1}^2 \to \overline{\delta}^2$. Hence under these conditions, I could use the magnitude

$$c_{t-1}^2 = \sum_{i=1}^{N_{t-2}} w_{i,t-2}\left(u_{i,t-1} - \frac{\lambda_i}{N_{t-2}w_{i,t-2}}u_{t-1}\right)^2 \tag{2.10}$$

directly for x_{t-1} in (2.4) to explore whether cross-sectional dispersion at time $t-1$ as summarized by the value of κ_{t-1}^2 contributes to aggregate market volatility σ_t^2.

Note that the above cross-sectional dispersion differs from cross-sectional market volatility of [69] in two ways[5]. First, I use individual forecasting errors instead of individual stock returns. More importantly, this formulation allows heterogeneous responses of individual stocks to the common market shock through the term λ_i and the lagged weight $w_{i,t-2}$.

One can fit a GARCH-X model to the aggregate return with the measure of cross-sectional dispersion given in (2.10). This provides a natural test of whether cross-sectional dispersion directly enters into the data-generating process for aggregate volatility.

---

[5]Two minor differences are 1) time-varying N_{t-2}, 2) predetermined individual weigths $w_{i,t-2}$ as opposed to $w_{i,t-1}$ in [69]. Section 2.4.2. provides robustness checks using cross-sectional market volatility of [69].

2.3 Results

In this section, I provide empirical evidence for the role of cross-sectional dispersion in predicting volatility of the S&P 500 Index. After describing the sample, I provide empirical evidence of improved specifications in GARCH-X models by allowing for feedback from cross-sectional dispersion to aggregate volatility, showing better in-sample forecasting performance. While GARCH-X with cross-sectional dispersion proves not to be robust for purposes of out-of-sample forecasts, I show that use of cross-sectional dispersion can help improve parameter estimates of a GARCH process for aggregate volatility by using individual stock returns jointly, generating better forecasts both in sample and out of sample.

2.3.1 Additional explanatory variable in GARCH

The dataset contains individual stocks comprising of S&P 500 Index from March 31, 1964 to December 31, 2014. The components of the S&P 500 Index have been updated periodically in response to corporate actions and index maintenance purposes for keeping the index up to date. The historical S&P 500 Index constituents are obtained from Compustat North America, and prices, cash dividends and the number of outstanding shares of individuals stocks are obtained from the Center for Research in Security Prices (CRSP) Security Files. There are total $1,545$ individual stocks that appear in the S&P 500 Index at least once during the sample period. Using individual stocks, I construct S&P 500 Index return from April 30, 1964 to December 31, 2014 both at a daily and a monthly frequency. While the S&P 500 Index adopted a free-float market capitalization methodology in 2005,[6] the S&P 500 Index return in this paper does not reflect such changes as it requires an additional adjustment in the number of shares used for calculating the index. Despite a minor deviation, the constructed index return is highly correlated

---

[6]S&P Dow Jones Indices gathers all public share ownership information and review each stock's Investable Weight Factor (IWF) that reflects only those shares available in the public markets on an annual basis. See more details about Investable Weight Factors and Float Adjustment Methodology from "S&P Dow Jones Indices: Index Methodology, January 2015".

with the quoted S&P 500 Index return; the correlation is 0.9968 prior to March 2005 and 0.9993 from April 2005 to December 2014. Furthermore, I also construct the S&P 500 Index return with dividends for studying the effects of ordinary cash dividends, which is used for robustness checks in section 2.4.3. The time period for the volatility forecasting exercises ranges from June 1964 ($t = 4$) to December 2014 ($t = 610$)[7] and the aggregate volatility of interest is equivalent to the volatility in monthly S&P 500 Index, that does not account for dividends.

To begin with, I describe the empirical procedure for constructing cross-sectional dispersion using individual stocks. First, for each individual stock i, I calculate a monthly return $r_{i,t}$ for $t = 2, \ldots, 610$, and estimate an individual forecasting error $u_{i,t}$ by regressing a stock i's return ($r_{i,t}$) on a constant and its lagged return ($r_{i,t-1}$) for $t = 3, \ldots, 610$. Second, a stock i's weight in period t, $w_{i,t}$ is calculated as $w_{i,t} = I_{i,t}p_{i,t}s_{i,t} / \sum_{j=1}^{N_t} I_{i,t}p_{j,t}s_{j,t}$, where $p_{i,t}$ is a price and $s_{i,t}$ is the number of outstanding shares of a stock i in period t for $t = 1, \ldots, 609$. Here $I_{i,t}$ is an indicator variable for the index constituents in period t, that is, $I_{i,t} \equiv 1$ if a stock i belongs to the S&P 500 index in period t and $I_{i,t} \equiv 0$ otherwise. Third, an aggregate forecasting error u_t is calculated from (2.8) given $u_{i,t}$ and $w_{i,t-1}$ for $t = 2, \ldots, 610$ and for all i. Next, I accommodate time-varying $\lambda_{i,t}$ in (2.6), allowing a stock i to respond to the aggregate market shock differently across time. This is intended to capture the nature of changes in the historical S&P 500 index constituents, of which periodic changes are mostly associated with corporate actions such as mergers, acquisitions and spin-offs, that are likely to change, changing firms' characteristics after such corporate actions. Time variation in firm-specific characteristics is also widely recognized in the empirical asset pricing literature. To do so, I repeat the first three steps using daily individual stocks, and obtain individual weights and forecasting errors as well as aggregate forecasting errors at a

---

[7]Monthly returns of individual stocks and accordingly constructed S&P 500 index are calculated from April 1964 to December 2014. Therefore, a monthly cross-sectional dispersion measure is constructed from May 1964 to November 2014, that is used for forecasting volatility of S&P 500 index from June 1964 to December 2014.

daily frequency. Then, for each stock i and month t, $\lambda_{i,t}$ is estimated by regressing $N_{t-1}w_{i,t-1}u_{i,t}$ on u_t using daily individual weights, forecasting errors and aggregate forecasting errors within each month t^8. Lastly, a measure for cross-sectional dispersion is provided by (2.10).

Figure 2.1 plots historical cross-sectional volatility, that is the square root of c_t^2 in (2.10). Notice that cross-sectional volatility measures average percentage deviation of individual forecasting errors from their aggregates, and reflects individual stocks' heterogeneity in two ways. First, it takes into account different degrees by which individual stocks respond to the aggregate shock provided through the term $\lambda_{i,t}$. Second, each stock's squared deviation contributes to the evolution of cross-sectional volatility proportional to its relative share in the S&P 500 Index given by $w_{i,t-1}$. The shaded areas represent seven NBER recession periods in the sample. Three points are worth noting. First, cross-sectional dispersion itself is time-varying and also exhibits high persistence with some clustering, as documented in [69] and [29]. Second, I find increasing cross-sectional dispersion during recession periods, indicating its potential role in helping to predict market volatility during recession periods. Lastly, cross-sectional dispersion is usually larger in magnitude than time-series market volatility (See Figure 2.2 for realized volatility), confirming the observation in [69]. The contemporaneous correlation between cross-sectional volatility and realized volatility is 0.4683, providing a rationale for considering the GARCH-X model with cross-sectional dispersion.

Given cross-sectional dispersion, I fit a GARCH-X model to the S&P 500 Index return as described by (2.2), (2.3) and (2.4). The second column of Ta-

---

[8]Specifically, let $d_t \in [1, \overline{m}_t]$ be the day of the month t, where $\max m_t = 23$ in the dataset. For each stock i and each month t, $\lambda_{i,t}$ is estimated by regressing $N_{t,d_t-1}w_{i,t,d_t-1}u_{i,t,d_t}$ on u_{t,d_t}, where N_{t,d_t}, w_{i,t,d_t}, u_{i,t,d_t} are number of stocks in the S&P 500 Index, stock i's weight and forecasting error at day d_t in month t. When a stock is newly added in the S&P 500 Index at the end of the month, $\lambda_{i,t}$ is replaced by its cross-sectional mean ($\lambda_{i,t} = 1$) within each month t. In general, the cross-sectional average of $\widehat{\lambda}_{i,t}$ is close to 1 with $N_t^{-1}\sum_{i=1}^{N_t}\widehat{\lambda}_{i,t} \in [0.98, 1.04]$. In order to guarantee the internal consistency, I normalize the cross-sectional average of $\widehat{\lambda}_{i,t}$ at each month t with $\widetilde{\lambda}_{i,t} = \widehat{\lambda}_{i,t}/\left(N_t^{-1}\sum_{i=1}^{N_t}\widehat{\lambda}_{i,t}\right)$.

ble 2.1 displays maximum likelihood estimates and asymptotic standard errors[9] for model parameters. For comparison, the third column reports those from a univariate GARCH(1,1) model with $\pi = 0$ in (2.4). In the numerical estimation procedure, the first observations (r_1 and c_1^2) are given and the initial value for aggregate volatility (σ_1^2) is jointly estimated with other model parameters although not reported.

While all parameter estimates for GARCH-X except ω and ϕ_1 are statistically significant at any conventional size, I find that the statistical significance for the coefficient estimate of cross-sectional dispersion (π) is inconclusive; given $H_0 : \pi = 0$, p-values from the Wald and the likelihood ratio tests[10] are respectively 0.15 and 0.05.

Using in-sample volatility forecasts implied by parameter estimates in Table 2.1, I evaluate forecasting performance by comparing forecasting accuracies of volatility forecasts from GARCH-X and GARCH models. I adopt realized variance as a proxy for latent volatility following [16], [64] and [83]. Denoting by $\sigma_{RV,t}^2$ realized variance in period t, it is measured by aggregating squared daily index returns within month t as,[11]

$$\sigma_{RV,t}^2 = \sum_{d=1}^{m_t} r_{d,t}^2 \tag{2.11}$$

where $r_{d,t}$ is a daily S&P 500 Index return at day d of month t and m_t is the number of trading days in month t.

Figure 2.2 plots realized volatility, which is a square root of realized variance in (2.11). Notice two periods with large observations; realized volatility in October 1987 is 26.63 and 23.94 in October 2008 while the historical average without these two observations is 3.9. The former is known as Black Monday when stock

---

[9]Asymptotic standard errors are estimated by approximating the second derivative of the log-likelihood functions at maximum likelihood estimates. See details for numerical MLE estimation and calculation of asymptotic standard errors in Hamilton [Time Series Analysis, 62, pp. 133-148].

[10]The Wald test statistic is $0.0306/0.0214 = 1.43$, and the 1 degree of freedom log-likelihood ratio test statistic is $2 \times \{-1,803.87 - (-1,805.81)\} = 3.8765$.

[11]For calculating realized variance, I splice the quoted daily S&P 500 Index return (prior to Feb 28, 2005) with the constructed daily S&P 500 Index return.

markets around the world crashed within a very short time and the latter is the peak of the recent Financial Crisis triggered by the U.S. housing bubble. These are widely recognized as outliers originating from a different distribution from the historical monthly realized volatility. Given their big impacts on volatility forecasting, I exclude volatility forecasts for October 1987 and those for four months from September 2008 to December 2008 when evaluating out-of-sample volatility forecasts later in this section.[12] Since observations for historical realized variance (or volatility) are widely used in the earlier literature,[13] I suppress further explanations and proceed to the evaluation of volatility forecasts using realized variance.

Table 2.2 reports average losses under three loss criteria; Mean Squared Error (MSE), Quasi-Likelihood (QLIKE) and Mean Absolute Error (MAE). Among various loss criteria, I consider MSE and QLIKE given their robustness to the presence of noise in the volatility proxy suggested by [83], and I also consider MAE given its robustness to a few influential forecasts. I provide definitions for each loss criteria in the second column. Columns 3 and 4 report average losses of in-sample volatility forecasts from GARCH and GARCH-X models. The adjacent row (in parenthesis) reports the percentage differences in forecasting accuracies of GARCH-X relative to GARCH, where for each loss function, the nagative difference implies that GARCH-X has smaller average losses than GARCH on average. I find that GARCH-X yields lower average losses under all loss criteria, and the largest improvement in forecasting accuracies is found under MAE loss fuction: GARCH-X yields 4.29% smaller average losses compared to the univariate GARCH model.

For investigating whether cross-sectional dispersion is useful for purposes of volatility forecasting in practice, I proceed to the evaluation of out-of-sample volatility forecasts. This is intended to address the potential over-fitting issue inso-

---

[12]In the presence of thorny outlier issues, [88] note that volatility models are generally ill-designed for predicting unforeseen and unprecedented extreme events.

[13]For example, [89] provide explanations for market volatility including definition, measurement and stylized facts about financial market volatility.

far as improved forecasting accuracy in GARCH-X may be provided by introducing an additional parameter in the data-generating process. I adopt a rolling fixed estimation period method[14] following [18]. With 5 years of estimation window, I fit models to a sample of 5 years, generate one-step ahead volatility forecasts and drop the oldest observation from the sample when adding the new data. I repeat this process and evaluate the performance of 542 monthly out-of-sample forecasts from June 1969 to December 2014[15]. Similarily, I also generate out-of-sample forecasts using 10 years of estimation window, evaluating 482 monthly forecasts from June 1974 to December 2014.

Table 2.3 compares out-of-sample forecasting accuracies of GARCH and GARCH-X models under 5 and 10 years of estimation window size. I provide definitions for each loss criteria in the second column. Columns 3 and 4 report average losses for out-of-sample volatility forecasts from GARCH and GARCH-X when one–step-ahead out-of-sample forecasts are obtained from parameter estimates using 5 years (Panel 1) and 10 years observations (Panel 2). As before, the adjacent row (in parenthesis) reports the percentage differences in forecasting accuracies of GARCH-X relative to univariate GARCH. Here, GARCH-X yields smaller average losses for all loss functions with 5 years of estimation window size, yet MSE from GARCH-X is slightly larger than univariate GARCH model when using 10 years of observations.

For statistical inference for average loss differentials, I perform tests for equal forecasting accuracy suggested by [34] and [101] (henceforth, DMW). De-

---

[14]Many researchers documented the merits of using a rolling fixed estimation period method among alternatives in addition to the ease of statistical inference. For example, [38], [54] and [18] noted that a rolling estimation period method is robust in the presence of nonstationarity. [102] showed that its forecasting accuracy is no worse than an expanding sample window method.

[15]For periods in October 1987 and those from September 2008 through December 2008, all models considered cannot forecast volatility accurately and generate extremely large losses across all loss functions. These in turn generate outliers in the empirical distribution of average loss differentials in DMW test statistic. For accurate statistical inference, I exclude five months' volatility forecasts among 547 monthly out-of-sample volatility forecasts from consideration. Though weakly statistically significant, results without excluding those volatility forecasts are generally consistent with those presented here. Table 2.10 compares forecasts with all out-of-sample volatility forecasts in parallel with Table 2.3.

noting by $loss_t$ the loss differential among competing forecasts in period t, an asymptotic pairwise test statistic for testing the null hypothesis of no difference in the forecasting accuracy is given by,

$$DMW = \frac{\overline{loss}}{avar\left(\overline{loss}\right)}$$

where $\overline{loss} = T^{-1} \sum_{t=1}^{T} loss_t$ is the sample mean loss differentials and $avar\left(\overline{loss}\right)$ is asymptotic variance of loss differentials. Following standard practice, I obtain a consistent estimate for $avar\left(\overline{loss}\right)$ by taking a weighted sum of the sample autocovariances using a Bartlett kernel.

For each loss function, the statistical significance of DMW test statistics is denoted by using an asterisk on the percentage differences in Table 2.3. In general, the statistical significance is inconclusive for GARCH-X. With 5 years of estimation window size, I find that average losses of GARCH-X are smaller than univariate GARCH for all loss functions and average loss differentials for QLIKE and MAE are statistically significant at 5%. With 10 years of estimation window size, however, none are statistically significant at any significance level while GARCH-X yields larger average MSE loss than univariate GARCH.

So far, I investigate one potential channel by which cross-sectional information might help predict volatility in the S&P 500 Index return. As an additional explanatory variable in GARCH process, cross-sectional dispersion helps predict volatility in sample across all loss criteria, and the statistically significant coefficient estimate of cross-sectional dispersion (π) based on the likelihood ratio test. Though improved out-of-sample forecasting accuracy in general, I find that the statistical significance of such improvements depends on the estimation window size, indicating the weak contribution of cross-sectional dispersion to the aggregate volatility generating process.

2.3.2 Aid in parameter estimation

Next, I investigate an alternative means by which cross-sectional dispersion might improve volatility forecasts. A convenient model for incorporating cross-sectional information is the factor-ARCH model developed in [44]. Although this model has been used in hundreds of studies, it has not been successfully applied to a cross-section of thousands of stocks due to computational difficulties[16]. In this section, I model a bivariate GARCH process for the aggregate volatility (σ_t^2) and the cross-sectional dispersion (κ_t^2), and estimate model parameters by jointly using a full cross-section of stock returns.

Using the same dataset containing prices and outstanding shares of stocks in monthly S&P 500 Index, I estimate parameters in the following model:

$$r_{i,t} = \phi_i^0 + \phi_i^1 r_{i,t-1} + u_{i,t}, \tag{2.12}$$

$$u_{i,t} = \frac{\lambda_i}{N_{t-1} w_{i,t-1}} \sigma_t \varepsilon_t + \frac{\kappa_t}{N_{t-1} w_{i,t-1}} \eta_{i,t}, \tag{2.13}$$

$$\varepsilon_t, \eta_{i,t} \sim \text{ i.i.d. } (0,1),$$

$$\begin{bmatrix} \sigma_t^2 \\ \kappa_t^2 \end{bmatrix} = \begin{bmatrix} \varpi_1 \\ \varpi_2 \end{bmatrix} + \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} u_{t-1}^2 \\ c_{t-1}^2 \end{bmatrix} + \begin{bmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{bmatrix} \begin{bmatrix} \sigma_{t-1}^2 \\ \kappa_{t-1}^2 \end{bmatrix} \tag{2.14}$$

where $\varpi_1, \varpi_2 > 0$ and $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_{11}, \beta_{22} \geq 0$ are model parameters. c_{t-1}^2 is a lagged cross-sectional dispersion measure provided in (2.10).

The aggregate volatility process in (2.14) is same as the GARCH-X model in (2.4) with $\varpi_1 = \varpi$, $\alpha_{11} = \alpha$, $\alpha_{12} = \pi$, $\beta_{11} = \beta$. Hence this suggests a statistical test for the direct role of cross-sectional dispersion in predicting aggregate volatil-

---

[16]To my knowlegde, the largest number of cross-sectional observations used within the factor-ARCH model is 50 in [42], where authors evaluate the performance of the class of covariance models including factor GARCH, restricted vector GARCH, dynamic conditional correlation GARCH models and extensions of these models.

ity by testing $H_0 : \alpha_{12} = 0$. Furthermore, the cross-sectional dispersion process is estimated along with the aggregate volatility process, where α_{12} and α_{21} capture the dynamic dependence between two volatility processes. Lastly, the above bivariate GARCH model includes Panel-ARCH provided by [25] as a special case: with $\lambda_i = 1$ for $\forall i$ and $\kappa_t = \tau$ for $\forall t$ in (2.13), Panel-ARCH estimates a univariate ARCH process for the aggregate volatility with $\alpha_{12} = \beta_{11} = 0$ in (2.14). In such a case, τ captures the time-series average of cross-sectional dispersion across a large number of cross-sectional observations. See also [25] for the estimation of quarterly profit uncertainty using industry-level sales revenues. Appendix 2.A describes the empirical procedure for estimating parameters in the above model.

Table 2.4 displays maximum likelihood estimates and asymptotic standard errors for model parameters. Parameter estimates from the bivariate GARCH process (2.14) are reported in the second column (Full Model). For comparison, the subsequent columns report parameter estimates from restricted models such as $\alpha_{21} = 0$ (Model 1) and $\alpha_{12} = \alpha_{21} = 0$ (Model 2). Parameter estimates for individual stocks such as ϕ_i^0 and ϕ_i^1, are not reported as they are obtained separately by a univariate regression for each i. All parameter estimates except ϖ_1 and α_{21} are statistically significant at any conventional size. Three points are worth noting. First, one-sided dynamic dependence is observed between the aggregate volatility and the cross-sectional dispersion processes. While confirming the direct contribution of cross-sectional dispersion in predicting market volatility from statistically significant α_{12}, [17] the coefficient estimate for α_{21} is statistically insignificant. This indicates the weak dynamic dependence between two volatility processes. Second, the cross-sectional volatility process is shown to be non-stationary in sample: the persistence of the process implied from coefficient estimates for α_{22} and β_{22} is 1.5981, which is greater than 1. This further contributes to the non-stationarity of the aggregate volatility process through α_{12}, yielding the coefficient estimate for ϖ_1

---

[17]P-value is 0.0087 from two-sided hypothesis test, and it is 0 from log-likelihood ratio test. The 1 degree of freedom test statistic is computed by comparing maximized log-likelihood values between Model 1 and Model 2, that is, $2 \times \{-1,328,838.98 - (-1,328,847.51)\} = 17.06$.

to be statistically insignificant. Lastly, the aggregate volatility process becomes to be stationary when shutting out the dynamic dependence from the cross-sectional volatility process. The aggregate volatility process in Model 2 is stationary with the persistence being 0.9606 which is close to the persistence of univariate GARCH model in Table 2.1, and with the long-run average being 14.68.

Next, I evaluate forecasting performance by comparing forecasting accuracies of volatility forecasts from univariate GARCH and the above bivariate GARCH models. More specifically, I compare forecasting accuracies of volatility forecasts from Model 1 and Model 2 with those from univariate GARCH. Though non-stationary in sample, I include Model 1 in the comparison of the forecasting performance in order to capture the possibility that cross-sectional dispersion could improve volatility forecasts out of sample when cross-sectional stock returns are jointly used for parameter estimation, which differs from the previous exercise using GARCH-X. Furthermore, it also enables us to infer the relative size of the direct contribution from the cross-sectional dispersion in Model 1 by comparing the forecasting performance of Model 1 with Model 2. It is because the improved forecasting accuracies of Model 2 relative to univariate GARCH can be viewed as being obtained indirectly by using cross-sectional stock returns jointly.

Results are provided at the last two columns in Table 2.2. There are two lines of empirical evidence supporting the improved in-sample forecasting accuracies by utilizing cross-sectional information. First, I find improved forecasting accuracies from both Model 1 and Model 2 across all loss criteria. In particular, the improved forecasting accuracies from Model 2 provides the evidence on the indirect contribution of cross-sectional information when cross-sectional stock returns are jointly used for estimating model parameters. Second, I find that Model 1 provides larger improvements in forecasting accuracies than Model 2, confirming the enhanced in-sample forecasting accuracies by directly using cross-sectional dispersion as an additional explanatory variable in the aggregate volatility process. Note that Model 1 provides smaller average losses across all loss criteria

compared to GARCH-X model. Under MAE loss criteria of which GARCH-X provides the largest improvement in forecasting accuracies by 4.29%, Model 1 has 30.22% smaller average losses than univariate GARCH, that is larger in magitude compared to GARCH-X.

Table 2.3 provides the evaluation of one-step-ahead out-of-sample forecasting accuracies from above two models when using 5 and 10 years of estimation windows. Again, I report the percentage differences relative to forecasts from GARCH as well as the statistical significance of DMW equal predictability test statistics using asterisks. Recall the inconclusive results from GARCH-X by failing to provide statistically significant improvements for out-of-sample volatility forecasts when using 10 years of estimation window. Here, I find the improved forecasting accuracies from the bivariate GARCH models. Model 2 has statistically significantly smaller average losses than GARCH under MSE and MAE loss functions at 5% significance level. When extending Model 2 by including cross-sectional dispersion in the aggregate volatility process, Model 1 has statistically significant improvements under MSE and MAE loss functions with 5 years of estimation window, and under all three loss functions with 10 years of estimation window. This contrasts with the failure of GARCH-X in predicting accurate volatility forecasts out of sample.

I further explore the improved forecasting abilities of the bivariate GARCH models by testing whether they also provides more accurate volatility forecasts in particular periods when accurate volatility forecasts are of great interest. More specifically, I test whether two models outperform univariate GARCH more during recessions when accurate volatility forecasts are of great interest. Let d_t be the loss differential between Model 1 and GARCH for predicting one-step-ahead out-of-sample volatility forecasts in period t. Using an indicator variable for NBER recession periods I_t^R, I perform tests for conditional predictive ability developed in [54] for testing $H_0 : E\left[d_t | \mathcal{F}_{t-1}\right] = 0$, which contrasts to $H_0 : E\left[d_t\right] = 0$ in DMW equal (unconditional) predictability tests.

Let h_t be a vector of variables that are thought to be important for relative

forecast performance; $h_t^R \equiv \left(1, d_t, I_t^R\right)'$ in this case. Given the conditional moment restriction, 3 degrees of freedom Wald-type test statistic (GW) is provided by,

$$GW = (T-1)\,\overline{Z}'\widehat{\Omega}^{-1}\overline{Z}$$

where $\overline{Z} \equiv (T-1)^{-1}\sum_{t=1}^{T-1} h_t^R d_{t+1}$ and $\widehat{\Omega} \equiv (T-1)^{-1}\sum_{t=1}^{T-1}\left(h_t^R d_{t+1}\right)\times\left(h_t^R d_{t+1}\right)'$ is a 3×3 matrix that consistently estimates the variance of $\left(h_t^R d_{t+1}\right)$. Under the null hypothesis, the test statistic is asymptotically chi-squared distributed with 3 degrees of freedom.

Panel 1 in Table 2.5 reports GW statistics using h_t^R with 5 years (columns 2-3) and 10 years of estimation windows (columns 4-5) respectively. During NBER recession periods, I find that two bivariate GARCH models provide more accurate out-of-sample volatility forecasts with 5 years of estimation window, that are statistically significant under all loss criteria (Model 1) and under two loss criteria except QLIKE (Model 2). With a larger estimation window such as 10 years, both models provide statistically significantly improved forecasts under all loss criteria. Panel 2 reports GW statistics for testing whether these models outperform univariate GARCH more during bear markets. I use another indicator variable I_t^N for periods with negative S&P 500 Index returns, and calculate GW statistics with $h_t^N \equiv \left(1, d_t, I_t^N\right)$. Results are similar to those using NBER recession periods. I find that above models yield more accurate forecasts during periods of negative stock returns, providing conditionally accurate volatility forecasts that are statistically significant under most loss criteria.

Although the forecasting equation is the same as univariate GARCH, I find the improved forecasting performance by jointly using cross-sectional information: volatility forecasts from Model 2 are more accurate than those from univariate GARCH both in sample and out of sample. Furthermore, I find additional improvements from Model 1 in predicting aggregate volatility when extending Model 2 by including cross-sectional dispersion in the aggregate volatility process.

The potential explanation for such improvement comes from the basic insight in [98] that when the number of cross-sectional observations is large, any aggregate factors can be uncovered essentially perfectly using the cross section. By jointly using the full cross section of stock returns, one can come up with better estimates of the population parameters. In other words, cross-sectional dispersion might help to estimate parameters in the aggregate volatility process.

To see this, consider the log-likelihood function of bivariate GARCH described by (2.12), (2.13) and (2.14). For expositional simplicity, I define alternative measures of the aggregate forecasting error and cross-sectional dispersion. Let $w_{i,t-1}^{\lambda} \equiv w_{i,t-1}\lambda_i$ with $w_{i,t-1}$ being a stock i's weight in (2.1) and λ_i being a stock i's firm-characteristic parameter in (2.6). The aggregate forecasting error in parallel with (2.8) becomes,

$$\widetilde{u}_t = \sum_{i=1}^{N_{t-1}} w_{i,t-1}^{\lambda} u_{i,t} \tag{2.15}$$

where $\widetilde{u}_t$ directly reflects firm-specific characteristics through different λ_is.

Denoting by $w_{i,t-1}^{w} \equiv N_{t-1}w_{i,t-1}^2$ another relevant weight of a stock i, the cross-sectional dispersion in parallel with (2.8) becomes,

$$\widetilde{c}_t^2 = \sum_{i=1}^{N_{t-1}} w_{i,t-1}^{w} \left(u_{i,t} - \frac{\lambda_i}{N_{t-1}w_{i,t-1}} u_t \right)^2 \tag{2.16}$$

where $\widetilde{c}_t^2$ assigns a higher weight to idiosyncratic forecasting errors of a stock with above-average weight $(w_{i,t-1} > 1/N_{t-1})$ compared to the cross-sectional dispersion in (2.8).

Using (2.15) and (2.16), the closed-form log-likelihood of bivariate GARCH

is given by,

$$
\mathscr{L} = \sum_{t=1}^{T} \left[-\frac{N_{t-1}}{2} \log 2\pi - \frac{1}{2} \log J_t \right. \tag{2.17}
$$

$$
\left. -\frac{N_{t-1}}{2} \left(\frac{1}{\kappa_t^2} \times \widetilde{c}_t^2 - \frac{\widetilde{\lambda}}{\kappa_t^2} \times \left(u_t - \frac{\widetilde{u}_t}{\widetilde{\lambda}} \right)^2 + \frac{\widetilde{\lambda}}{\kappa_t^2 + \sigma_t^2 \cdot N_{t-1}\widetilde{\lambda}} \times \frac{\widetilde{u}_t^2}{\widetilde{\lambda}^2} \right) \right]
$$

where $J_t \equiv \left(\kappa_t^2 + \sigma_t^2 \sum_{i=1}^{N_{t-1}} \lambda_i^2 \right) \times (\kappa_t^2)^{N_{t-1}-1} \times \prod_{i=1}^{N_{t-1}} (w_{i,t-1} N_{t-1})^{-2}$ is a determinant of a $N_{t-1} \times N_{t-1}$ variance matrix of individual forecasting errors, and $\widetilde{\lambda} \equiv N_{t-1}^{-1} \sum_{i=1}^{N_{t-1}} \lambda_i^2$.

The above log-likelihood function (2.17) shows that cross-sectional information enters in the log-likelihood function through two channels. Cross-sectional dispersion ($\widetilde{c}_t^2$), weighted by $1/\kappa_t^2$, helps squared errors (u_t^2) to estimate model parameters. Furthermore, heterogeneous characteristics of stock market index constituents (λ_is) feed further cross-sectional information into the estimation procedure through $\widetilde{u}_t$ and $\widetilde{\lambda}$, where distances between u_t and $\widetilde{u}_t/\widetilde{\lambda}$ become to be momentous. These provide the rationale behind the improved forecasting accuracies of Model 2 relative to univariate GARCH although not using cross-sectional dispersion directly for predicting aggregate volatility.

Next, I separate heterogeneous λ_is' contribution from the improved forecasting abilities observed in bivariate GARCH models. To do so, I consider the simple case when all stocks have same magnitudes in responding to the aggregate market shock ($\lambda_i = 1$ for $\forall i$), resulting in $\widetilde{\lambda} = 1$ and $\widetilde{u}_t = u_t$. Then, the closed-form log-likelihood function reduces to,

$$
\mathscr{L} = \sum_{t=1}^{T} \left[-\frac{N_{t-1}}{2} \log 2\pi - \frac{1}{2} \log \overline{J}_t - \frac{N_{t-1}}{2} \left(\frac{1}{\kappa_t^2} \times \overline{c}_t^2 + \frac{1}{\kappa_t^2 + \sigma_t^2 \cdot N_{t-1}} \times u_t^2 \right) \right] \tag{2.18}
$$

where $\overline{c}_t^2$ is calculated by imposing $\lambda_i = 1$ for $\forall i$ in (2.16) and $\overline{J}_t$ is a determinant associated with individual forecasting errors in (2.6) with $\lambda_i = 1$ for $\forall i$.

Table 2.6 compares forecasting accuracies of in-sample volatility forecasts

from GARCH-X and two bivariate GARCH models when $\lambda_i = 1$ for $\forall i$. For comparison, the first column repeat the average losses from univariate GARCH displayed in Table 2.2. The preceeding columns report the percentage differences in forecasting accuracies of GARCH-X, Model 1 and Model 2 relative to univariate GARCH. When excluding heterogeneous λ_is from the volatility forecasting models, I find from Panel 1 that in-sample volatility forecasts, in general, are less accurate than those reported in Table 2.2. Strikingly, GARCH-X become to lose its improvements in forecasting accuracies across all loss functions, that is mostly due to the missing information previously provided by λ_is. In contrast, Model 1 remains to have improved volatility forecasts relative to univariate GARCH although exhibiting smaller improvements across all loss criteria when comparing with Table 2.2. I further compare forecasting accuracies of out-of-sample volatility forecasts when using 5 years (Panel 2) and 10 years (Panel 3) of estimation window, yet again confirming the enhanced forecasting accuracies from two bivariate GARCH models by jointly using a full cross-section of stock returns.

Lastly, I provide the log-likelihood of univariate GARCH, which is a special case of (2.17). Assuming $\lambda_i = 1$ for $\forall i$ and $\kappa_t^2 = \kappa^2$, $c_t^2 = c^2$ for $\forall t$, the maximization of the above log-likelihood is equivalent to,

$$\mathscr{L} = \sum_{t=1}^{T} \left[-\frac{1}{2} \log 2\pi - \frac{1}{2} \log \sigma_t^2 - \frac{1}{2} \times \frac{u_t^2}{\sigma_t^2} \right] \tag{2.19}$$

where only squared forecasting errors are used for parameter estimation.

2.4   Robustness checks

In this section, I provide robustness checks for the empirical analysis provided in Table 2.2 and Table 2.3. First, I consider squared return as a proxy for unobservable volatility. Second, I address concerns for highly dispersed cross-sectional returns and corresponding cross-sectional kurtosis by considering alter-

native cross-sectional dispersion measures. Third, I take ordinary cash dividends into account and evaluate the forecasting performance across models in predicting volatility of the S&P 500 Index return with dividends. Lastly, I present the volatility forecasting performance when outliers are included for the evaluation of forecasting accuracies. Thoughout this section, I provide the comparison of volatility forecasting accuracies across models, where Panel 1 compares in-sample forecasting accuracies and Panel 2 (Panel 3) compares out-of-sample forecasting accuracies when using 5 years (10 years) estimation window. For each panel in the tables, the second column reports average losses from GARCH while the preceeding three columns report the percentage differences in forecasting accuracies relative to univariate GARCH.

2.4.1 Alternative volatility proxy

As a proxy for latent volatility, squared returns have been widely adopted in the earlier literature (See, for example, [82], [31] and [51]). As an alternative volatility proxy, I construct a monthly squared return proxy by squaring residuals[18] from the AR(1) forecasting model in (2.2). Table 2.7 provides results for volatility forecasting performance when squared returns are used for comparing forecasting accuracies across models. Results are consistent with those reported in Table 2.2 and Table 2.3.

2.4.2 Alternative measures

In this subsection, I provide robustness checks by using alternative cross-sectional dispersion measures within the GARCH-X model. This comes from the recognition in [69] that cross-sectional returns are highly dispersed, especially for a large number of stocks considered. For investigating whether empirical results are

---

[18]Denoting by $\sigma^2_{SR,t}$ squared return in period t, it is calculated by $\sigma^2_{SR,t} = \left(r_t - \widehat{\phi}^0 - \widehat{\phi}^1 r_{t-1} \right)^2$ where $\widehat{\phi}^0$ and $\widehat{\phi}^1$ are OLS estimates obtained by regressing a monthly return on a constant and its lag.

affected by cross-sectional kurtosis, I consider three alternative measures: 1) 5% trimmed estimator, 2) squared interquartile range and 3) cross-sectional market volatility suggested in [69]. Let $y_{i,t-1} \equiv w_{i,t-2}^{1/2} (u_{i,t-1} - \lambda_{i,t-1} u_{t-1})$ for $t = 2, \ldots, T$. Rewriting the equation for cross-sectional dispersion in (2.10) as,

$$c_{t-1}^2 = \sum_{i=1}^{N_{t-1}} y_{i,t-1}^2$$

Then, the 5% trimmed estimator is obtained from the above equation by discarding 5% extreme $y_{i,t-1}$ from both tails at each period $t - 1$. Similarily, the squared interquartile range estimator is calculated by squaring interquartile range of $y_{i,t-1}$ at each period $t - 1$.

The cross-sectional market volatility is constructed as[19],

$$\sigma_{C,mt-1}^2 = \sum_{i=1}^{N_{t-1}} w_{i,t-2} (r_{i,t-1} - r_{t-1})^2$$

Figure 2.3 plots cross-sectional volatility from three alternative measures, displayed as a square root of cross-sectional dispersion. Note that magnitudes of the first two alternative measures are smaller than Figure 2.1, and the cross-sectional market volatility exhibits similar patterns to the baseline measure in Figure 2.1[20].

I estimate a GARCH-X model again, this time replacing c_{t-1}^2 in (2.10) by the above three alternatives. Table 2.8 compares forecasting accuracies of GARCH-X using three alternative measures both in sample and out of sample. In general, results are consistent with those reported in Table 2.2 and Table 2.3. All three alternative measures provide more accurate in-sample volatility forecasts than uni-

---

[19]From the cross-sectional market volatility in [69], I made two adjustments for reflecting the time-varying number of stocks (N_{t-1}), and for making weights of individual stocks to be predetermined $(w_{i,t-2})$ in period $t - 1$. While the former adjustment is crucial in the analysis, I find that the effect from the latter adjustment is trivial.

[20]One exception is found during the stock market crash in October 1987, where the cross-sectional market volatility fails to capture the hightened idiosyncratic volatility documented in [29].

variate GARCH model[21], however, the improved forecasting accuracies are not statistically significant out of sample in general[22].

Before concluding this section, it is worth noting that the cross-sectional market volatility provides more accurate out-of-sample volatility forecasts than a univariate GARCH model. This finding contrasts to [69] of which volatility forecasts of individual stocks are shown to be marginally improved out of sample[23]. However, such improvements in forecasting accuracies within the GARCH-X model are smaller than from the cross-sectional dispersion in this paper both in sample and out of sample (See Table 2.2, 2.3 and 2.8). This is mainly due to the implicit assumption in [69]'s measure; all individual stocks are treated as if they respond to the aggregate and idiosyncratic shocks in a same manner. However, it is less likely to be a valid assumption when considering individual stocks' heterogeneities originated from disparities in market capitalizations as well as differences in firm-specific characteristics.

2.4.3   Ordinary cash dividends

While ordinary cash dividends account for significant variations in daily individual stock returns, the S&P 500 Index return is generally quoted without dividends, capturing the changes in the index constituents' prices only. In this subsection, I evaluate the forecasting performance across models in predicting the volatility of the S&P 500 Index return with dividends. In other words, the aggregate volatility of interest is changed to the volatility in monthly S&P 500 Index

---

[21]The coefficient estimate of cross-sectional dispersion (π) is statistically significant at 5% (10%) when using 5% trimmed estimator and cross-sectional market volatility measure.

[22]Two exceptions are under MAE loss criteria when using the trimmed cross-sectional dispersion measure, and under QLIKE loss criteria when using [69]'s cross-sectional market volatility measure. Both are statistically significant at the 10% level.

[23]Potential explanations for this are associated with 1) an easiness of forecasting volatility of the aggregate market index rather than of an individual stock, and 2) an exhaustive consideration of the index constituents rather than excluding some stocks during the sample period . I also consider [69]'s GARCH-cross-sectional (GARCH-XC) model by excluding the constant coefficient from GARCH-X in (2.4) as $\sigma_t^2 = \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 + \pi x_{t-1}$. Though not reported, GARCH-XC does not provide statistically significant improvements in forecasting accuracies out of sample.

return with cash dividends. To do so, I construct a fictional S&P 500 Index return with dividends where ordinary cash dividends are treated as if they are reinvested, and estimate GARCH, GARCH-X and two bivariate GARCH models using monthly S&P 500 Index return with dividends. Table 2.9 compares forecasting accuracies across models both in sample and out of sample. Results are consistent with those reported in Table 2.2 and Table 2.3. This confirms the improved forecasting accuracies of bivariate GARCH models provided by using cross-sectional information, which is also applicable for predicting the aggregate volatility in the stock index return with dividends.

2.4.4 Outlier issues

So far, I exclude out-of-sample volatility forecasts for two periods (October 1987 and Septermber-December 2008) when comparing forecasting accuracies across models. Since stock market volatility during these periods is unforeseen and unprecedented, all forecasting models based on historical price information under-forecast volatility of the S&P 500 Index return, commonly resulting in large losses across all loss criteria regardless a proxy for unobserved volatility. In this subsection, I evaluate forecasting models' performance again when previously excluded forecasts are included for calculating out-of-sample forecasting accuracies.

Table 2.10 reports the comparison of out-of-sample forecasting accuracies of univariate GARCH, GARCH-X and two bivariate GARCH models when using 5 years (Panel 1) and 10 years of estimation window (Panel 2). While results, in general, are consistent with those presented in Table 2.3, I find the effects of thorny outliers on forecasting accuracies in two ways. First, all forecasting models generate severely under-predicted volatility forecasts during two influential periods. GARCH-X and two bivariate GARCH models have smaller improvements in forecasting accuracies (relative to univariate GARCH) compared to those in Table 2.3. In particular, they become to have larger average QLIKE losses than univariate GARCH (See Panel 1), where QLIKE loss function imposes heavier penalty

on under-prediction. Second, severely under-forecasted volatilities during these periods also affect the statistical significance of DMW test statistics. When using MSE loss function, Model 1's forecasting accuracies are not different from univariate GARCH's in statistical sense. Weakened statistical significance associated with MSE loss function is mainly due to incorrect asymptotic variance estimates, that are heavily affected by outliers. On the other hand, Model 1 and Model 2 remain to have statistically smaller average MAE losses than univariate GARCH, indicating smaller influences from outliers when using MAE loss function.

2.5   Conclusion

This paper investigates the role of cross-sectional information in predicting aggregate volatility. Given a large number of individual stocks, I develop a model of stock returns by reflecting a natural idea that individual stocks respond to the common aggregate shock at different degrees. The model is simple, but it also provides a natural measure for cross-sectional dispersion whose effects on the stock market volatility and cyclical variations in macroeconomic variables have been popular research topics.

Using individual stocks in the S&P 500 Index from March 1964 to December 2014, I test the direct contribution of cross-sectional dispersion in predicting stock market volatility. Although helpful for in-sample volatility forecasts, GARCH-X with cross-sectional dispersion fails to provide more accurate out-of-sample volatility forecasts than GARCH. In other words, I provide empirical evidence that cross-sectional dispersion does not enter the data-generating process for market volatility.

I further explore another possibility for cross-sectional dispersion contributing to accurate estimates for model parameters. Using full cross-section of individual stocks jointly, I estimate parameters in the bivariate GARCH model of aggregate volatility and cross-sectional dispersion. I find that the cross-sectional

dispersion improves the accuracies of aggregate volatility forecasts both in sample and out of sample. Furthermore, out-of-sample volatility forecasts from the bivariate GARCH model are shown to be more accurate than those from GARCH in times of NBER recessions as well as during the periods with negative S&P Index returns.

Given improved forecasting accuracies when using cross-sectional stock returns jointly, I conclude that cross-sectional dispersion does help predict volatility forecasts indirectly by helping to estimate parameters. On the other hand, empirical evidence from GARCH-X indicates that it does not enter in the data-generating process directly.

2.6   Acknowledgements

Chapter 2, in full, has been submitted for publication of material.

Figures and Tables
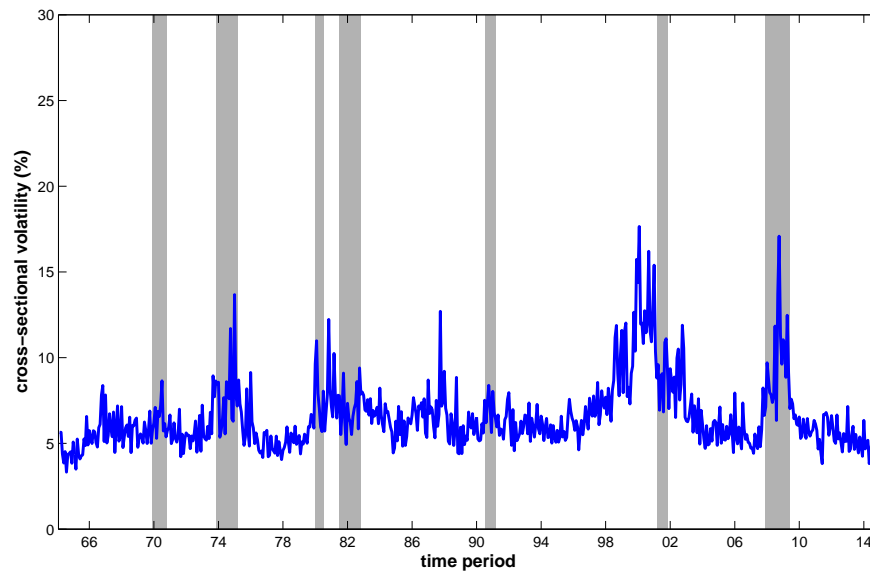
Figure 2.1: Cross-sectional dispersion

Figure 2.1 plots historical cross-sectional volatility which is a square root of cross-sectional dispersion across individual stocks following (2.10). Shaded areas represent NBER recession periods. The cross-sectional dispersion is shown to be time-varying and highly persistent. The cross-sectional dispersion spiked up during the stock market crash in October 1987, a feature commonly observed from alternative measures considered in earlier literature.
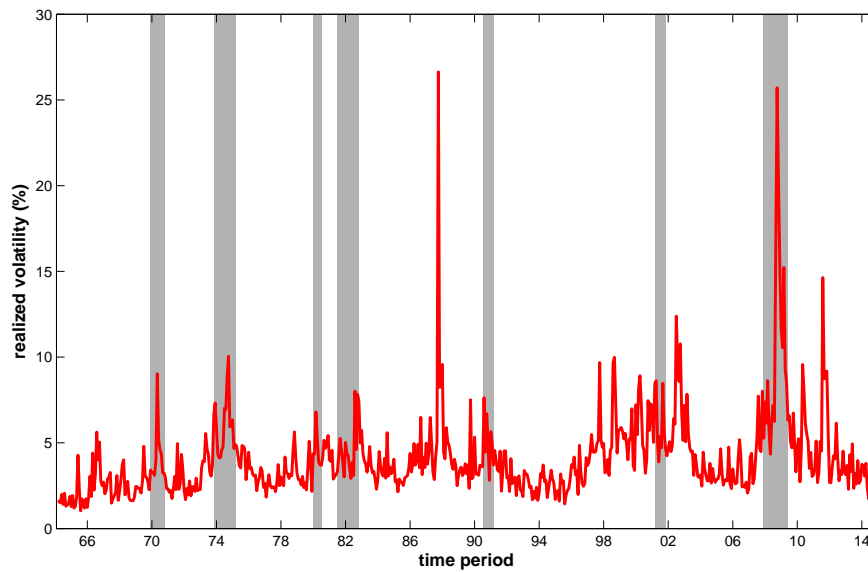
Figure 2.2: Realized volatility

Figure 2.2 plots historical realized volatility which is a square root of realized variance calculated by aggregating daily S&P 500 Index returns following (2.11). Given a free-float market capitalization methodology adopted in 2005, the daily S&P 500 Index returns are calculated using individual stock returns and their weights, and I splice the quoted daily S&P 500 Index return (prior to Feb 28, 2005) with the constructed daily S&P 500 Index return for calculating realized variance. Shaded areas represent NBER recession periods.
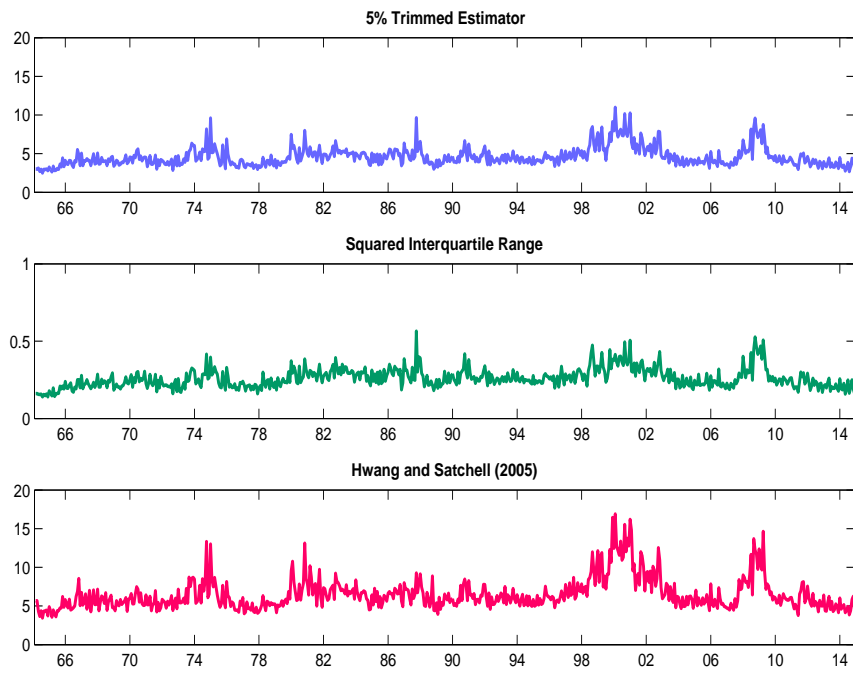
Figure 2.3: Alternative cross-sectional dispersion measure

Figure 2.3 plots alternative cross-sectional dispersion measures. Top panel displays cross-sectional volatility after removing 5% extreme observations from both tails. Middle panel displays interquartile range in (2.10). Bottom panel plots cross-sectional market volatility proposed by Hwang and Satchell (2005). The plotted are the square root of alternative cross-sectional dispersion measures.

Table 2.1: Parameter estimates

Parameters	GARCH-X		GARCH	
	MLE	(s.e.)	MLE	(s.e.)
ϖ	0.4303	(0.5174)	0.9713	(0.3503)
α	0.1032	(0.0289)	0.1176	(0.0285)
β	0.8295	(0.0354)	0.8510	(0.0266)
π	0.0306	(0.0214)		
ϕ_0	0.8712	(0.1755)	0.8724	(0.1735)
ϕ_1	0.0085	(0.0442)	0.0106	(0.0441)
Likelihood	$-1,803.87$		$-1,805.81$	

Table 2.1 reports MLE estimates (asymptotic standard errors) of model parameters in GARCH-X and GARCH models. Asymptotic standard errors are estimated by approximating the second derivative of the log-likelihood functions at MLE estimates. The last row reports the maximized log-likelihood values under two volatility forecasting models.

Table 2.2: Comparison of in-sample forecasting accuracy

$Criteria$	$L\left(\sigma_{RV}^2, \sigma^2\right)$	GARCH	GARCH-X	Model 1	Model 2
MSE	$(\sigma_{RV}^2 - \sigma^2)^2$	$1,916.32$	$1,860.34$	$1,826.23$	$1,871.44$
			(-2.92)	(-4.70)	(-2.34)
$QLIKE$	$\frac{\sigma_{RV}^2}{\sigma^2} + \log \sigma^2$	3.96	3.94	3.92	3.95
			(-0.54)	(-1.16)	(-0.33)
MAE	$\lvert \sigma_{RV}^2 - \sigma^2 \rvert$	17.43	16.68	12.16	13.23
			(-4.29)	(-30.22)	(-24.09)

Table 2.2 provides the comparison of in-sample volatility forecasts from univariate GARCH, GARCH-X and two bivariate GARCH models provided in Section 3.2.. As a proxy for an unobservable volatility, historical realized variance is calculated from the daily S&P 500 index return. Column 2 provides definitions of loss functions for measuring volatility forecasting accuracies. Next three columns report average losses of in-sample volatility forecasts from univariate GARCH, GARCH-X models and two bivariate GARCH models. Adjacent rows report the percentage differences in forecasting accuracies of GARCH-X, bivariate GARCH models relative to GARCH, where the negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH.

Table 2.3: Comparison of out-of-sample forecasting accuracy

Panel 1 : 5 years of estimation window					
Criteria	$L\left(\sigma_{RV}^2, \sigma^2\right)$	GARCH	GARCH-X	Model 1	Model 2
MSE	$(\sigma_{RV}^2 - \sigma^2)^2$	711.17	593.45	428.26	436.82
			(-16.55)	(-39.78^{**})	(-38.58^{**})
QLIKE	$\frac{\sigma_{RV}^2}{\sigma^2} + \log\sigma^2$	4.15	4.09	4.14	4.18
			(-1.25^{**})	(-0.19)	(0.72)
MAE	$\|\sigma_{RV}^2 - \sigma^2\|$	17.91	16.34	11.61	12.57
			(-8.74^{**})	(-35.16^{***})	(-29.83^{***})
Panel 2 : 10 years of estimation window					
Criteria	$L\left(\sigma_{RV}^2, \sigma^2\right)$	GARCH	GARCH-X	Model 1	Model 2
MSE	$(\sigma_{RV}^2 - \sigma^2)^2$	567.18	567.25	419.68	398.19
			(0.10)	(-26.01^{**})	(-29.79^{***})
QLIKE	$\frac{\sigma_{RV}^2}{\sigma^2} + \log\sigma^2$	4.18	4.13	4.11	4.14
			(-1.06)	(-1.67^{**})	(-0.99)
MAE	$\|\sigma_{RV}^2 - \sigma^2\|$	17.33	16.36	11.52	11.97
			(-5.54)	(-33.50^{***})	(-30.92^{***})

Table 2.3 provides the comparison of out-of-sample volatility forecasts from univariate GARCH, GARCH-X and two bivariate GARCH models provided in Section 3.2.. A rolling fixed estimation period method was used for calculating out-of-sample volatility forecasts with 5 years (Panel 1) and 10 years of the estimation window (Panel 2). Column 2 provides definitions of loss functions for measuring volatility forecasting accuracies. Next three columns report average losses of in-sample volatility forecasts from univariate GARCH, GARCH-X models and two bivariate GARCH models. Adjacent rows report the percentage differences in forecasting accuracies of GARCH-X, bivariate GARCH models relative to GARCH. While the negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH, asterisk represents the statistical significance of DMW equal predictability test suggested by [34] and [101]. Given critical values being 1.28 (90%), 1.65 (95%) and 2.33 (99%) respectively, */**/*** represent the statistical significance at 90%, 95% and 99% respectively.

Table 2.4: Parameter estimates

Parameters	Full Model		Model 1		Model 2	
	MLE	(s.e.)	MLE	(s.e.)	MLE	(s.e.)
ϖ_1	0.0000	(0.4505)	0.0000	(0.4505)	0.5785	(0.2552)
ϖ_2	12.7429	(0.4640)	12.7430	(0.4640)	12.7429	(0.4640)
α_{11}	0.0819	(0.0269)	0.0819	(0.0269)	0.1163	(0.0262)
α_{12}	0.0428	(0.0180)	0.0428	(0.0180)	−	−
α_{21}	0.0000	(0.0111)	−	−	−	−
α_{22}	0.9893	(0.0194)	0.9893	(0.0189)	0.9893	(0.0189)
β_{11}	0.7922	(0.0534)	0.7922	(0.0534)	0.8443	(0.0285)
β_{22}	0.6088	(0.0075)	0.6088	(0.0075)	0.6088	(0.0075)
Likelihood	$-1,328,838.98$		$-1,328,838.98$		$-1,328,847.51$	

Table 2.4 reports MLE estimates (asymptotic standard errors) of the bivariate GARCH model. Asymptotic standard errors are estimated by approximating the second derivative of the log-likelihood functions at MLE estimates. For comparison, it also reports estimation results under two restricted models: $\alpha_{21} = 0$ (Model 1) and $\alpha_{12} = \alpha_{21} = 0$ (Model 2). The last row reports the maximized log-likelihood values under three models.

Table 2.5: Tests for conditional predictive ability

Panel 1 : Recession				
	5 years		10 years	
Criteria	Model 1	Model 2	Model 1	Model 2
MSE	39.02***	34.25***	34.85***	58.29***
QLIKE	26.56***	–	46.50***	43.67***
MAE	96.35***	76.77***	104.27***	123.25***
Panel 2 : Bear markets				
	5 years		10 years	
Criteria	Model 1	Model 2	Model 1	Model 2
MSE	36.31***	26.77***	41.57***	59.92***
QLIKE	26.78***	–	47.13***	43.54***
MAE	107.00***	85.53***	119.29***	137.19***

Table 2.5 reports test statistics for conditional predictive ability proposed by [54]. Under the null hypothesis of no conditional loss differentials, the test statistic is asymptotically chi-squared distributed with 3 degrees of freedom. Panel A report results for testing whether two bivariate models outperform univariate GARCH during recession with using 5 years and 10 years estimation windows. Panel B reports results for testing whether those models outperform univariate GARCH during periods with negative S&P 500 Index return. Given critical values with 3 degrees of freedom being 6.25 (90%), 7.81 (95%) and 11.34 (99%) respectively, */**/*** represent the statistical significance at 90%, 95% and 99% respectively.

Table 2.6: Comparison of forecasting accuracies with all $\lambda_i = 1$

Panel 1 : In-sample				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	$1,916.32$	6.70	-0.34	-2.58
QLIKE	3.96	0.47	-0.11	-0.61
MAE	17.43	3.93	-20.02	-21.64
Panel 2 : Out-of-sample (5 years of the estimation window)				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	711.17	54.55	-38.11***	-38.41**
QLIKE	4.15	-0.68	-0.60	-0.32
MAE	17.91	0.45	-30.30***	-26.76***
Panel 3 : Out-of-sample (10 years of the estimation window)				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	567.18	60.87	-25.72***	-26.94***
QLIKE	4.18	0.31	-0.28	-0.60
MAE	17.33	0.40	-29.18***	-26.96***

Table 2.6 highlights the indirect contribution of the cross-sectional dispersion for volatility forecasts. Compared to previous forecasting exercises, univariate GARCH-X and two bivariate models are estimated with strict restrictions on firm-characteristics: all $\lambda_i = 1$. Panel 1 compares in-sample forecasts from univariate GARCH, GARCH-X and two bivariate GARCH models, and Panel 2 (Panel 3) compare their out-of-sample volatility forecasts with 5 years (10 years) of the estimation window. While column 2 reports average losses of volatility forecasts from GARCH, the preceeding three columns report the percentage differences in forecasting accuracies of GARCH-X and two bivariate GARCH models relative to univariate GARCH model. The negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH, and asterisk represents the statistical significance of DMW equal predictability test suggested by [34] and [101].

Table 2.7: Rubustness checks - Squared return

Panel 1 : In-sample				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	$1,461.98$	-1.58	-8.22	-6.74
QLIKE	3.89	-0.26	-0.82	-0.43
MAE	22.95	-1.70	-21.18	-18.42
Panel 2 : Out-of-sample (5 years of the estimation window)				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	$1,299.22$	-11.18	-31.33**	-28.92**
QLIKE	4.03	-1.12**	-1.03	-0.27
MAE	24.79	-4.57*	-25.38***	-21.90***
Panel 3 : Out-of-sample (10 years of the estimation window)				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	$1,105.39$	-3.67	-17.32***	-17.47***
QLIKE	4.01	-1.09*	-1.99**	-1.39*
MAE	23.78	-3.11	-21.48***	-20.89***

Table 2.7 provides the comparison of forecasting accuracies across models when using squared return as a proxy for volatility forecasts. Monthly squared returns are constructed by squaring residuals from the AR(1) forecasting model for the S&P 500 Index return. For comparison, column 2 reports average losses of volatility forecasts from GARCH and the preceeding three columns report the percentage differences in forecasting accuracies of GARCH-X and two bivariate GARCH models relative to univariate GARCH model. The negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH, and asterisk represents the statistical significance of DMW equal predictability test suggested by [34] and [101].

Table 2.8: Robustness checks - Alternative measures

Panel 1 : In-sample				
Criteria	GARCH	5 %	IQR	HS
MSE	$1,916.32$	-3.09	-2.73	-2.39
QLIKE	3.96	-0.55	-0.44	-0.47
MAE	17.43	-4.47	-3.54	-3.47
Panel 2 : Out-of-sample (5 years of the estimation window)				
Criteria	GARCH	5 %	IQR	HS
MSE	711.17	-3.48	16.56	-8.82
QLIKE	4.14	-0.58	0.13	-0.78^*
MAE	17.90	-6.26^*	0.16	-5.39
Panel 3 : Out-of-sample (10 years of the estimation window)				
Criteria	GARCH	5 %	IQR	HS
MSE	567.18	-2.38	-1.76	3.84
QLIKE	4.18	-0.44	-0.61	-1.00
MAE	17.33	-3.69	-4.58	-4.09

Table 2.8 reports robustness checks using three alternative cross-sectional dispersion measures. Given concerns for a noisy cross-sectional dispersion measure when using all individual stock returns, I consider two alternative measures as a covariate in the GARCH-X model; in (2.10), I calculate 1) the trimmed cross-sectional dispersion (5%) where 5% extreme observations are removed from both tails at each t, 2) squared interquartile range (IQR). Lastly, the cross-sectional market volatility is calculated following Hwang and Satchell (2005). For comparison, column 2 reports average losses of volatility forecasts from GARCH and the preceeding three columns report the percentage differences in forecasting accuracies of GARCH-X using three alternative measures relative to univariate GARCH model. The negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH, and asterisk represents the statistical significance of DMW equal predictability test suggested by [34] and [101].

Table 2.9: Robustness checks - Dividends

Panel 1 : In-sample				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	$1,795.00$	-2.76	-5.64	-2.51
QLIKE	3.97	-0.47	-1.28	-0.46
MAE	17.33	-3.75	-29.47	-23.10
Panel 2 : Out-of-sample (5 years of the estimation window)				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	710.09	-13.83	-43.62***	-41.80***
QLIKE	4.16	-0.86*	-0.33	0.58
MAE	18.07	-6.95*	-35.33***	-30.36***
Panel 3 : Out-of-sample (10 years of the estimation window)				
Criteria	GARCH	GARCH-X	Model 1	Model 2
MSE	559.98	1.04	-24.39**	-24.74***
QLIKE	4.19	-0.85	-1.35**	-0.54
MAE	17.02	-3.38	-27.62***	-24.60***

Table 2.9 provides the comparison of volatility forecasts from univariate GARCH, GARCH-X and two bivariate GARCH models provided in Section 3.2.. For taking into account the effects of ordinary cash dividends on volatility forecasting accuracy, I constructed S&P 500 index with dividends and compare volatility forecasting accuracies across models. As a proxy for an unobservable volatility, historical realized variance is calculated from daily returns for the constructed S&P 500 index with dividends. For comparison, column 2 reports average losses of volatility forecasts from GARCH and the preceeding three columns report the percentage differences in forecasting accuracies of GARCH-X and two bivariate GARCH models relative to univariate GARCH model. The negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH, and asterisk represents the statistical significance of DMW equal predictability test suggested by [34] and [101].

Table 2.10: Robustness checks - Outliers

Panel 1 : Out-of-sample (5 years of the estimation window)				
$Criteria$	GARCH	GARCH-X	Model 1	Model 2
MSE	$2,284.72$	-2.56	-6.21	-5.34
$QLIKE$	4.22	0.11	0.74	1.98
MAE	20.96	-7.23^{**}	-28.43^{***}	-23.85^{***}
Panel 2 : Out-of-sample (10 years of the estimation window)				
$Criteria$	GARCH	GARCH-X	Model 1	Model 2
MSE	$2,333.14$	3.48	-1.18	-5.39^{**}
$QLIKE$	4.23	-0.22	-0.52	-0.26
MAE	20.76	-3.95	-26.53^{***}	-25.24^{***}

Table 2.10 provides the comparison of out-of-sample volatility forecasts from univariate GARCH, GARCH-X and two bivariate GARCH models provided in Section 3.2. without excluding out-of-sample forecasts from two influential periods. Panel 1 (Panel 2) compares out-of-sample volatility forecasts from univariate GARCH, GARCH-X and two bivariate GARCH models with 5 years (10 years) of the estimation window. While column 2 reports average losses of volatility forecasts from GARCH, the preceeding three columns report the percentage differences in forecasting accuracies of GARCH-X and two bivariate GARCH models relative to univariate GARCH model. The negative difference implies that corresponding forecasting model has smaller average losses than univariate GARCH, and asterisk represents the statistical significance of DMW equal predictability test suggested by [34] and [101].

Appendix 2.A: Empirical Procedure

In this section, I describe the parameter estimation procedure of the proposed model which is skipped in Section 2.3.2. For clarifying dimensions associated with vectors and matrices, I use one underline below a variable for representing a vector and two underlines for representing a matrix.

To begin with, consider individual forecasting errors obtained from AR(1) forecasting model of returns (2.12): for each stock i, an individual forecasting error $u_{i,t}$ is obtained by regressing $r_{i,t}$ on a constant and its lagged return $r_{i,t-1}$ for $t = 2, \ldots, T$. Let $\underline{u}_t \equiv [u_{1,t}, \ldots, u_{N_t,t}]'$ be a collection of individual forecasting errors at period t. Denoting by $\underline{\underline{\Omega}}_t \equiv E\left[\underline{u}_t \underline{u}_t' | \mathcal{F}_{t-1}\right]$ a $N_t \times N_t$ variance-covariance matrix of individual forecasting errors in period t, the joint log-likelihood function of individual stock returns becomes,

$$\mathscr{L} = \sum_{t=1}^{T} \left[-\frac{N_{t-1}}{2} \log 2\pi - \frac{1}{2} \log \left| \underline{\underline{\Omega}}_t \right| - \frac{1}{2} \underline{u}_t' \underline{\underline{\Omega}}_t^{-1} \underline{u}_t \right]$$

In general, the numerical maximization of above log-likelihood by iterative methods can be quite costly since it requires an inversion and a determinant calculation of $N_t \times N_t$ matrix $\underline{\underline{\Omega}}_t$ for each period t.

Here I overcome this empirical intractability issue by modeling an individual forecasting error using a factor structure provided by equation (2.13). Since $\underline{\underline{\Omega}}_t$ is a symmetric matrix that is factored by a vector of individual weights, analytical forms of inversion and determinant are given by,

$$\left| \underline{\underline{\Omega}}_t \right| = \left(\kappa_t^2 \right)^{(N_{t-1}-1)} \cdot \left(\kappa_t^2 + \sigma_t^2 \sum_{i=1}^{N_{t-1}} \lambda_i^2 \right) \cdot \left(\prod_{i=1}^{N_{t-1}} \frac{1}{N_{t-1}^2 w_{i,t-1}^2} \right)$$

$$\underline{\underline{\Omega}}_t^{-1}(i,j) = \begin{cases} \frac{N_{t-1}^2 w_{i,t-1}^2}{\kappa_t^2} \cdot \left(1 - \frac{\sigma_t^2 \lambda_i^2}{\kappa_t^2 + \sigma_t^2 \sum_{i=1}^{N_{t-1}} \lambda_i^2} \right) & \text{for } j = i \\ -\frac{N_{t-1}^2 w_{i,t-1} w_{j,t-1}}{\kappa_t^2} \cdot \frac{\sigma_t^2 \lambda_i \lambda_j}{\kappa_t^2 + \sigma_t^2 \sum_{i=1}^{N_{t-1}} \lambda_i^2} & \text{for } j \neq i \end{cases}$$

where $\underline{\underline{\Omega}}_t^{-1}(i,j)$ is an $(i,j)^{th}$ element in the inverse matrix $\underline{\underline{\Omega}}_t^{-1}$.

Then, the closed-form log-likelihood function is given by,

$$
\begin{aligned}
\mathscr{L} \;=\; \sum_{t=1}^{T} \Bigg[& -\frac{N_{t-1}}{2} \log 2\pi - \frac{1}{2} \log J_t \\
& - \frac{N_{t-1}^2}{2\kappa_t^2} \left\{ \sum_{i=1}^{N_t} \left(w_{i,t-1} u_{i,t} \right)^2 - \frac{\sigma_t^2}{\kappa_t^2 + \sigma_t^2 \sum_{i=1}^{N_{t-1}} \lambda_i^2} \left(\sum_{i=1}^{N_t} w_{i,t-1} \lambda_i u_{i,t} \right)^2 \right\} \Bigg]
\end{aligned}
$$

where $J_t \equiv \left(\kappa_t^2 + \sigma_t^2 \sum_{i=1}^{N_{t-1}} \lambda_i^2 \right) \times (\kappa_t^2)^{(N_{t-1}-1)} \times \prod_{i=1}^{N_{t-1}} (w_{i,t-1} N_{t-1})^{-2}$ is a determinant of a $N_{t-1} \times N_{t-1}$ variance matrix of individual forecasting errors. This can be numerically evaluated along with the bivariate GARCH process, providing MLE estimates for parameters.

Chapter 3

Heterogeneity in the Dynamic Effects of Uncertainty on Investment

Abstract. We examine how aggregate profit uncertainty influences capital investment activities, focusing on heterogeneous responses of firms. We model aggregate profit uncertainty as the conditional standard deviation of a common factor across unforecasted fluctuations in the sales growth of different industries, and exploit cross-sectional variations for its estimation. From an investment forecasting model that coherently analyzes firm- or group-specific effects of uncertainty, we find that the direction and the size of investment adjustment vary considerably across firms, with a significant but small negative average impact. Our results highlight the importance of accounting for heterogeneity in the transmission of uncertainty, allowing us to reconcile different views on the effect of uncertainty in the existing literature.

3.1   Introduction

In the seminal work of Dixit and Pindyck [36], uncertainty is a key factor influencing investment decisions of firms, and a large body of literature has been subsequently devoted to study this topic. Nonetheless, earlier discussions arrive at divergent conclusions about the uncertainty-investment relationship. For example, Hartman [65] and Abel [1] find that a mean-preserving increase in uncertainty raises capital investments under convex adjustment costs. In a single-project partial equilibrium model, Sarkar [93] shows the positive impact of an uncertainty increase below a threshold. More recently, Slade [97] finds empirical evidence of a positive relation, using data for the US mining sector. On the contrary, [9] and Bloom [11] points out the irreversibility of capital, which causes firms to wait under high uncertainty. Baum et al. [8]'s empirical analysis supports this view. Our goal in this paper is to emphasize the importance of accounting for the heterogeneous responses of firms to changes in aggregate profit uncertainty, to better understand the opposing views in the previous literature.

Considering heterogeneity by exploring the firm-level data essentially opens up the possibility of clearly understanding various features at work that different theoretical models focus on. To this end, we present an investment forecasting model where both the sign and the magnitude of investment adjustment are coherently examined at the firm level. This model differs from the approach widely employed in previous studies on the differential effects of uncertainty (e.g., Leahy and Whited 79 and Ghosal and Loungani 53). These studies first select a certain firm feature as a potential cause of the heterogeneity and split the sample accordingly into two groups for analysis. This may not explicitly capture the differences in the uncertainty transmission mechanism across sub-samples, as other explanatory variables can also have varying effects simultaneously on investment. In our framework, however, only profit uncertainty is interacted with firm-specific slope coefficients, allowing for a straightforward investigation of the transmission

channel. We first let the data speak with regard to firm-level heterogeneity, and examine ex-post how various firm characteristics contribute to the heterogeneity. Furthermore, our framework can be easily modified to accommodate a number of firm features and grouping criteria to examine potential causes of heterogeneity. It is also widely applicable to a number of other contexts, for instance, to investigate how asset holdings or loan growth of financial intermediaries change at the firm level when aggregate economic uncertainty rises.

The second contribution of the paper is to propose a volatility process called Panel-autoregressive conditional heteroskedasticity (Panel-ARCH) to efficiently estimate aggregate profit uncertainty from a large panel of industries. We model aggregate profit uncertainty as the conditional standard deviation of a common factor that simultaneously drives unforecasted fluctuations in sales growth across different industries. In this setting, we postulate factor loadings as well as idiosyncratic volatilities to be inversely proportional to the previous quarter's sales share of each industry. This enables the derivation of a closed-form log-likelihood function that not only expedites the estimation process but also provides more precise parameter estimates for the aggregate volatility process. Thus, Panel-ARCH builds on the factor-ARCH models developed in Diebold and Nerlove [35] and [44], as it jointly models individual entities using a factor structure. Yet our approach further simplifies the computational process, making the model easily applicable for an aggregate volatility estimation from a panel with a large cross-sectional dimension.

We find that profit uncertainty on average has a significant but small negative effect on investment. However, when we look at the firm-level responses, substantial heterogeneity is observed across firms: about 28% of firms in the sample are expected to increase capital investment when uncertainty is high, and the magnitude of investment adjustment varies considerably across firms. Thus, our finding allows reconciling the different results in the previous literature, and highlights the importance of accounting for heterogeneity to better understand the

transmission of uncertainty. Further, when we link the observed heterogeneity to firm size, we find that the size and the degree of uncertainty effects have an inverse U-shape relationship. That is, both small and large firms are expected to reduce investment more than medium-sized firms. This non-linear effect of uncertainty with respect to the firm size has not been documented in the previous literature such as Ghosal and Loungani [53] where the relationship is demonstrated to be rather linear based on a two-group analysis. On the contrary, our findings corroborate the theoretical result illustrated in Bolton et al. [14] that the relation between a firm's available internal funds and investment timing is highly non-monotonic and non-linear, assuming that the size represents the firm's financial condition. We further evaluate several firm characteristics that are relevant to the heterogeneity, such as the total-liabilities-to-asset ratio, and the sub-industry classification.

The remainder of the paper is as follows. The next section presents the Panel-ARCH process for aggregate profit uncertainty. Section 3.3 introduces an investment capital decision forecasting model capturing heterogeneous effects of uncertainty, and briefly describes the data. The estimation results are presented in section 3.4 including a detailed discussion of different firm characteristics and heterogeneity in the uncertainty transmission mechanism. Section 3.5 concludes.

3.2 Measuring Profit Uncertainty

3.2.1 Profit Uncertainty

A firm's profitability is determined by several unobservable factors such as consumers' taste, wealth, production technology, and prices of inputs and outputs. Nonetheless, we use a firm's sales revenue as a proxy of profitability since it is a fundamental source of the firm's periodic profit. Sales revenue (henceforth sales) is also less likely to be affected by the firm's non-operating activities such

as profits/losses from its subsidiary as well as its financial investment activities.[1]

We define aggregate profit uncertainty as the conditional standard deviation of a factor which simultaneously drives unexpected changes in the sales growth rates across different industries.[2] Support for such a common driver is reported in Herskovic et al. [67], documenting a strong factor structure in volatilities of firms' sales growth. Another link between aggregate and industry-level sales can be found from the following accounting identity. For a total of J industries' sales (S_t^j for $j = \{1, ..., J\}$) in the economy, an aggregate sales index (S_t) is defined as

$$S_t \equiv \sum_{j=1}^{J} S_t^j, \tag{3.1}$$

and its growth rate from t to $t+1$ can be calculated as

$$g_{t+1} \equiv \frac{S_{t+1} - S_t}{S_t} = \sum_{j=1}^{J} w_{j,t} \cdot g_{t+1}^j. \tag{3.2}$$

where $w_{j,t} \equiv S_t^j / \sum_{j=1}^{J} S_t^j$ is the sales share of an industry j in period t and g_{t+1}^j is the growth rate of the industry $j's$ sales from t to $t+1$. Hence, the aggregate sales growth rate is the weighted average of the industries' sales growth rates, where the weights are time-varying and pre-determined as the t period's sales share.[3] It is worthwhile to note that the weights have a further implication regarding how much variations in each industry contribute to the aggregate-level volatility. Therefore, we model aggregate profit uncertainty as the variability of a common factor across the sales forecast errors of different industries.

We use the sales series of all firms in the combined quarterly Compustat

---

[1]Ghosal and Loungani [53] also constructed an annual measure of profit uncertainty from the residual of a sales forecasting equation of large and small firms, defined as 5-year standard deviations of the residuals.

[2]Our classification of industries in this paper is based on the first 2-digit Standard Industrial Classification (SIC) code. The same classification is used in [57] when estimating the marginal profitability of capital using a sales-revenue-to-capital ratio.

[3]Appendix 3.A.2 provides plots of the quarterly aggregate sales revenue index following (3.1) as well as the quarterly aggregate sales growth rate following (3.2).

North America industrial files from 1981Q1 to 2012Q4. We construct a panel of 67 2-digit standard-industry-classification (SIC) industries by aggregating sales within the same industry, and calculate quarterly sales growth rates as well as sales shares.[4] For estimating unforeseen changes of industries, we consider a one-period-ahead sales growth forecasting model for an industry j, controlling for observable macroeconomic conditions Z_t (i.e., the quarterly real GDP growth rate and effective federal funds rate) and seasonality D_t as

$$g_{t+1}^{j} = \delta^{j} g_t^{j} + \phi^{j\prime} Z_t + \psi^{j\prime} D_t + u_{j,t+1}, \tag{3.3}$$

where coefficients δ^j, ϕ^j and ψ^j are allowed to vary across different industries. Then, based on a set of cross-sectional forecasting errors $u_{j,t+1}$, we propose a Panel-ARCH model for aggregate profit uncertainty in the following section.

3.2.2 Panel-ARCH Model

We exploit a factor model framework to capture aggregate profit uncertainty as the conditional standard deviation of a common factor driving cross-sectional sales forecasting errors of industries. More specifically, with f_{t+1} denoting the factor, we conjecture that industry $j's$ forecasting error (u_{t+1}^{j}) has the following factor structure:

$$u_{j,t+1} = \lambda_{j,t} \cdot (f_{t+1} + \eta_{j,t+1}), \tag{3.4}$$

where $\eta_{j,t+1}$ is industry $j's$ idiosyncratic forecasting error following a martingale difference sequence with $E\left(\eta_{j,t+1}^2 | \mathcal{F}_t\right) = \tau^2$, with an information set $\mathcal{F}_t$ containing all information available through t. We define aggregate profit uncertainty, σ_{t+1},

---

[4]There are 74 industries based on 2-digit SIC code. In our analysis, we exclude 7 industries whose sales revenues are infinitesimal or do not exist on a continuation basis over the sample period. Those are Agricultural Product-Livestock & Animal Specialties (SIC2=2), Forestry (SIC2=8), Fishing, Hunting and Trapping (SIC2=9), and other miscellaneous services (SIC2=81,84,86,89). Their sales revenues range from 0.00001% to 0.00308%, and account for 0.0063% as a whole from 1981Q1 to 2012Q4.

as the conditional standard deviation of f_{t+1}, i.e.:

$$E\left[f_{t+1}^2|\mathcal{F}_t\right] = \sigma_{t+1}^2, \tag{3.5}$$

provided that $E\left[f_{t+1}|\mathcal{F}_t\right] = 0$.

The factor structure in (3.4) is similar in spirit to the factor-ARCH models proposed in Diebold and Nerlove [35] and Engle et al. [44] in jointly modeling individual entities. However, the flexible multivariate specification of common and idiosyncratic volatilities in their framework made the application of their model to a large panel difficult, as the number of parameters to estimate increases rapidly along with the size of a cross-sectional dimension. Hence, we propose a Panel-ARCH model which improves on this feature by imposing a structure on $\lambda_{j,t}$ as following:

$$\lambda_{j,t} = \frac{1}{J \cdot w_{j,t}}. \tag{3.6}$$

This restriction implies that the volatility of each industry differs from others depending on its sales share in the previous period ($w_{j,t}$). More specifically, industry j with a below-average share (i.e., $w_{j,t} < 1/J$) tends to have large unpredicted variations, since the small industry is likely (i) affected heavily by the common uncertainty and (ii) under large idiosyncratic fluctuations. Figure 3.1 shows that the correlation coefficients between the absolute value of forecasting errors ($|u_{j,t+1}|$) and sales shares ($w_{j,t}$) across industries in each quarter are indeed negative for most periods in the sample, i.e., 113 out of 123 quarters, supporting the above specification. Furthermore, it ensures that the conditional variance of the weighted sum of forecasting errors across industries is equal to the conditional variance of the common factor as

$$E\left[\left(\sum_{j=1}^{J} w_{j,t} \cdot u_{j,t+1}\right)^2 |\mathcal{F}_t\right] = \sigma_{t+1}^2. \tag{3.7}$$

Therefore, aggregate profit uncertainty can be estimated from a recursive

formulation as follows:

$$
\begin{aligned}
\sigma_{t+1}^2 &= a_0 + \sum_{k=1}^{q} a_k \cdot u_{t-k+1}^2 \\
&= a_0 + \sum_{k=1}^{q} a_k \cdot \left(\sum_{j=1}^{J} w_{j,t-k} u_{j,t-k+1} \right)^2,
\end{aligned} \tag{3.8}
$$

where $u_{t-k+1} \equiv \sum_{j=1}^{J} w_{j,t-k} u_{j,t-k+1}$ is the weighted sum of forecasting errors across industries.

In Appendix 3.A.1, we demonstrate that a simple closed-form expression for the likelihood function can be derived under condition (3.6), in addition to the recursive formulation of σ_{t+1}^2. Thus, the parsimonious specifications of the Panel-ARCH provides tractability for the estimation of aggregate profit uncertainty, overcoming empirical difficulties due to a large cross-section. In addition, [24] documents the improved accuracy in aggregate volatility forecasting in a similar context: he shows that the cross-sectional information of stock index components results in more accurate parameter estimates and further leads to the more precise estimation of stock index volatility.

Here we fit a Panel-ARCH(1) for aggregate profit uncertainty from 1982Q1 to 2012Q4 (124 quarters).[5] Figure 3.2 plots the estimated profit uncertainty process with four shaded NBER recession periods. The uncertainty series shows frequent fluctuations over time, even though we first remove seasonality from the sales series by including seasonal dummy variables for each industry, when estimating the sales forecasting model (3.3).[6] The profit uncertainty was high during

---

[5]We select the Panel-ARCH(1) after comparing it to a few other specifications. The parameter estimates of the model imply that the profit uncertainty has persistence (α_1) of 0.2085 and the long-run average ($E\left[\sigma_t^2\right]$) of 12.24. Estimated τ indicates that average cross-sectional dispersion is approximately 33.67%. For model parameter estimation, we condition on the first 4 quarter observations (i.e., 1981Q1 - 1981Q4). More details of the comparison results as well as the parameter estimates are in Appendix 3.A.3.

[6]When we regress the estimated profit uncertainty series on four seasonal dummies, we reject the F-test that all seasonal dummies are jointly zero. However, to show that the results of this paper are not driven by the potential seasonality of the uncertainty series, we estimate all models again using a de-seasonalized profit uncertainty series, and the results do not change.

the 1981 and 2001 recessions, and then surged to an unprecedented level during the Great Recession. We also see jumps corresponding to the Mexican peso crisis in the mid-1990s as well as the Asian and Russian financial Crisis in the late-1990s. Although the size of an increase is relatively small, it also picked up during the episode of Black Wednesday in 1992 when Britain left the European Exchange Rate Mechanism.

3.3 An Investment Forecasting Model

In this section, we first describe our firm-level data with control variables for explaining the investment activity of individual firms. Next, we propose an investment forecasting model where firms' heterogeneous adjustment of investment under uncertainty is examined with estimated one-quarter-ahead aggregate profit uncertainty (denoted as $\sigma_{t+1|t}$).

We use the manufacturing firms (SIC codes 2000-3999) in the combined quarterly Compustat North America industrial files from 1989Q1 to 2012Q4.[7] Since only a few firms span the entire sample period while most emerge (or disappear) in the midst of this period, we construct an unbalanced panel after deleting observations that are missing, highly distressed or likely to be of mergers. The final sample contains 219,538 firm-quarter observations for 96 quarters and 5,197 firms.

Our main variable of interest is the future investment-to-capital-stock ratio of a manufacturing firm $y_{i,t+1}$. In the investment forecasting model, we control for the firm's investment opportunity, internal funding ability and life-cycle behavior, using sales revenue (sales), cash and cash-equivalent stocks (cash), and the logarithm of total book-valued assets (size).[8] These controls are expected to explain

---

[7]We focus on the manufacturing industry among others, since the capital expenditures of manufacturing firms take the largest share of total capital expenditure, and thus understanding their variations is critical. According to the US Census Bureau's data, the manufacturing industry in the U.S. comprised about 19.7% of total capital expenditures in 2000. The share fell to 15.4% in 2009; however, manufacturing still remains the largest contributor to total capital expenditures.

[8]While the average Tobin's Q had been a popular proxy since [66] for a firm's investment

the firm's future profitability as well as its capital investment decision. We follow the standard approach in the literature of transforming trends in variables (i.e., capital expenditure, sales and cash) into ratios relative to the beginning-period capital stocks. Lastly, we adjust the total book-valued assets to 2005Q3 US dollars. Table 3.1 reports the summary statistics along with the definitions of the selected variables for the analysis. For instance, the quarterly investment rate (y_{it+1}) has a mean of 6.3% with an inter-quantile range of 4.7%. In general, all variables exhibit a high degree of kurtosis, with the exception of size. More details on the data is provided in Appendix 3.B.

Next, we describe our investment forecasting model, where the profit uncertainty is interacted with firm-specific slope coefficients. Let $y_{i,t+1}$ be the investment-to-capital-stock ratio of a manufacturing firm i for $i = 1, ..., n$. Let $X_{i,t}$ be a 5×1 vector of firm-specific determinants of the future investment: following the investment literature (e.g., Gilchrist and Himmelberg 57), we include sales, cash holdings, size, a constant and lagged investment in $X_{i,t}$. We assume a linear forecasting model:

$$y_{i,t+1} = X'_{i,t} \cdot \beta + \gamma_i \cdot \sigma_{t+1|t} + \delta \cdot Z_t + h_{i,t+1}, \tag{3.9}$$

where β is a 5×1 vector of coefficients that are common across all firms and $h_{i,t+1}$ is an i.i.d. idiosyncratic forecasting error with variance ζ^2. To make sure that our results are not driven by the business cycle property of uncertainty, we also include the real GDP growth rate (Z_t) to control for cyclical variations in the first moment.

The coefficient of our main interest is γ_i, which is unique for each firm i: it quantifies the heterogeneous effects of profit uncertainty on the next period's investment activity. When γ_i is set to be the same across firms, it simply examines the average effect of uncertainty, with the proposed model devolving to a standard

---

opportunity, we use the sales revenue instead, following [57]. Since there has been a long-standing consensus regarding the inappropriateness of the average Tobin's Q, earlier Q-based investment regressions were often augmented by various measures of cash flows. See, for example, the literature reviews in [28] and [68].

forecasting model. Hence, the proposed framework can provide a detailed picture of changes at the firm level in a coherent manner, with γ_i varying alone for each firm. Furthermore, we can easily modify the model to accommodate pre-selected grouping criteria with group-specific slope coefficients.

We use three different model specifications to understand the uncertainty transmission mechanism. First, we estimate the above investment forecasting model ignoring heterogeneity, assuming identical γ across all i's (i.e., $\gamma_i = \bar{\gamma} \ \ \forall \ i$). In what follows, we call this the baseline analysis. Second, we estimate the model with complete firm-level heterogeneity (i.e., $\gamma_i \neq \gamma_j \ \ \forall \ i \neq j$). Every firm in the sample has its own γ_i, implying that each can respond differently to aggregate profit uncertainty. In this specification, we estimate the model by iterating between a linear projection of γ_i's and the numerical maximum likelihood estimation (MLE) of other parameters recursively, until convergence. This handles an empirical difficulty arising from the large dimension of γ_i, adding more than 5,000 firm-specific parameters to the estimation. Finally, we group firms based on their common characteristics such as firm size, total liability ratio, and the 2-digit SIC classification, and illustrate differences in responses by including group-specific slope coefficients (i.e., $\gamma_i = \bar{\gamma}_J \ \ \forall \ i \in J$). In this way, characteristics potentially related to the heterogeneity are highlighted, linking results closely to economic fundamentals.

3.4 Results

3.4.1 Findings from the Baseline Model

Profit Uncertainty (γ)

Column (1) of Table 3.2 reports coefficient estimates and asymptotic standard errors of the baseline model *without* incorporating heterogeneity.[9] In the

---

[9]We estimate the asymptotic standard errors by numerically approximating second derivatives of the Hessian matrix at the estimated MLE. See pp.133-148 in [62] for more details.

baseline model, firms' future capital investments are, on average, slightly nega-
tively related to our measure of aggregate profit uncertainty: a one-unit increase
in uncertainty will decrease firms' investment-to-capital ratios by 0.139 pp across
all firms, i.e., about 2.21% of the average quarterly investment growth rate (6.3%).
When the profit uncertainty peaked in 2008Q4, our result implies that firms' in-
vestment decreased by 0.99 pp due to a one-time surge in the uncertainty from
4.98% (2008Q3) to 8.87% (2008Q4). This result of the average negative effect of
uncertainty is consistent with the findings in Bloom [11]. Yet the size of the un-
certainty effect is relatively small, which emphasizes the importance of examining
the underlying transmission channel more closely at the firm level.

Other Firm-Specific Controls

Sales (β_1)

Sales have been used as a proxy of investment opportunities in the existing
literature, with the advantage of being available for both private and public firms
(see Acharya et al. 2, for example). We find that sales positively affect capital
investment, in line with previous studies such as [6]: assuming that sales represent
a firm's marginal productivity of capital and profit opportunity, it is more likely for
a firm with high sales to make capital investment. More specifically, when the sales-
to-capital ratio rises by one standard deviation (219.9%), investment increases by
1.34 pp, about 20% of both the size of its mean and standard deviation.

Cash (β_2)

Cash holdings are an internal source of funding, which can be particularly
helpful when a firm is financially constrained, as noted in [50], [86] and [33]. Con-
firming the previous findings, the model estimates that the cash holdings of a firm
are positively related to the investment forecast: firms with high cash holdings
have large internal funds supporting future investments. Additionally, if the cash
holdings of a firm reflect profitability, firms with more cash are more profitable,
and hence invest more.

Size (β_3)

Our baseline specification result shows that the coefficient of size is positive, but not significant. This may arise due to complex dynamics underlying size. First, from the perspective of a firm's life-cycle behavior, young and thus smaller firms are expected to be more profitable and invest more. However, size represents many other characteristics than just a firm's life cycle. For example, [57] consider it to be informative of a firm's financial condition, since it may represent the stability of its on-going business activity and/or the degree of public information available for investors. Larger firms are also likely to have easier access to external financing.[10] What we find is likely to be the result based on the mixture of various driving forces.

3.4.2 Heterogeneity at the Firm Level

Given the average significant but small negative effect of profit uncertainty, we now examine heterogeneity across firms in the uncertainty transmission mechanism by estimating the most flexible version of the model, where all firms are allowed to react differently to uncertainty. Since firm-specific slope coefficients (γ_i's) may vary across individual firms, this gives heterogeneity the best chance to be influential. Coefficient estimates for other control variables are similar to those in the baseline case (see Column (2) in Table 3.2) except for that of size which became significantly positive, implying that larger firms are likely to invest more.

The estimated firm-specific slope coefficients indicates significant heterogeneity in the transmission channel of profit uncertainty. In order to highlight the firm-level heterogeneity, Figure 3.3 plots a histogram of γ_i estimates across 5,197 firms. The histogram shows substantial heterogeneity, although the average of the γ_i's is negative.[11] About 28% of firms in our sample would respond positively to

---

[10]Firm size is hence frequently used to split samples, to distinguish between financially constrained and unconstrained firms (see Gertler and Gilchrist 52 and Carpenter et al. 27, among others).

[11]The average of the point estimates of γ_i's is -0.323 and the median is -0.256, consistent

profit uncertainty by increasing their capital expenditures. Moreover, among firms that are expected to decrease their future capital expenditure, some are much more severely affected by the uncertainty than others.

To better illustrate the observed heterogeneity, the frequency of γ_i's in each bin is stacked by firm-size quintiles in different colors in Figure 3.3. It is clear that the profit uncertainty is transmitted into firms' capital investment decisions by different magnitudes, conditional on their sizes. For instance, a large proportion of firms in the third and fourth quintiles (Medium and Large groups) expand capital expenditures under high profit uncertainty. More importantly, the observed relationship between the effect of uncertainty and firm size is not monotonic. Many firms in the largest as well as smallest size groups are also hit harshly and contract investment activities, which has not been observed in the previous literature such as Ghosal and Loungani [53], who show that the negative impact of profit uncertainty is substantially greater in industries dominated by small firms, implying a linear monotonic relationship between the two.

Next, we stack the bins by quintiles based on firms' total-liability-to-asset ratios in Figure 3.4. As anticipated, we see that the effect of profit uncertainty is more severe for firms with higher liability ratios. At the same time, a fare share of firms in the first and second quintiles are still expected to increase investment. Related to this point, [15] show that it can be beneficial for a firm with low future liquidity to make an early investment if it considers waiting too risky, as the delayed investment also pertains to elevated uncertainty in future funding availability.

In sum, our analysis highlights that firms differ considerably in the way they adjust investment activity when aggregate uncertainty fluctuates. The result indicates that various features of the firms studied in previous theoretical literature on the transmission of uncertainty, such as the irreversibility of capital and capital with time to build, are simultaneously at work by different degrees, emphasizing the importance of examining disaggregated level data to better understand the

---

with the baseline estimate of γ (-0.139) presented in section 3.4.1.

transmission mechanism.

3.4.3   Firm Characteristics and Uncertainty

Here we first group firms based on size, liability ratio, and 2-digit SIC sub-industry classification, and include corresponding group-specific coefficients of uncertainty in the investment forecasting model. While this specification lets heterogeneity be manifested to a lesser degree, it presents results in a concise way. Plus, the estimation is done by MLE alone in this case, affording a chance to evaluate the model formally using likelihood ratio test statistics.

The last three columns in Table 3.2 report the coefficient estimates for firm-specific controls, which are similar to the preceding two columns. Hence we next discuss in detail how different firm characteristics are related to the heterogeneity, abstaining from other firm-level controls. Table 3.3 reports estimated parameters for group-level heterogeneity together with brief descriptions of the grouping criteria.

Size

We first divide firms by size, allocating them into five groups based on their average total assets during the sample period.[12] Panel 1 of Table 3.3 reports the number of firms, and the average size in each group. It also presents the estimates for group-specific coefficients of uncertainty, illustrated again in Figure 3.5.

Our main finding here is that the relationship between the size and the magnitude of response is not linear, but has an inverse U-shape: in fact, firms of the largest size group respond as negatively as those of the smallest size group to profit uncertainty, simplifying the result shown in Figure 3.3. It is worth noting, again, that the inverse U relationship between the size and the magnitude of response has not previously been found in previous studies, mainly because of a framework that

---

[12]The grouping is based on the firms' average total assets, which are book-valued, denominated in 2005Q3 U.S. dollars and averaged within their presence during the sample periods.

admits a small number of groups. For example, [41] documents that small firms' timings of investments are more sensitive to profits compared to those of large firms, and Ghosal and Loungani [53] find that smaller firms are affected more negatively by profit uncertainty. Compared to the earlier approaches, our analysis can accommodate a larger number of groups, easily visualizing any underlying non-linearity in the relationship. The inverse U relationship is still robust when we increase the number of the size groups.

As noted earlier, size has often been thought of as a proxy for financial constraints. In that regards, we find evidence corroborating the theoretical result illustrated recently in Bolton et al. [14] where the relation between the level of internal funds and investment timing of a firm is rather non-monotonic and non-linear. In addition, larger firms are likely to have higher capital adjustment costs, and thus have higher option values of waiting. Together, these features can result in firms in the largest group being less resilient to uncertainty changes.

Liability Ratio

The leverage ratio (total-assets-to-liabilities ratio) of firms has been extensively used to identify firms with external funding abilities (see Bernanke et al. 10 for an extensive literature review). For example, [78] note that high leverage reduces a firm's financing ability to pursue a profitable investment opportunity through a liquidity effect. Using a total-liabilities-to-assets ratio (hereafter, liability ratio) as a proxy for a firm's financial condition, we re-estimate the group-specific slope coefficients of uncertainty.

Panel 2 of Table 3.3 reports the number of firms and the average liability ratio in each group, followed by γ_i's estimates; the results are plotted in Figure 3.6. The results show that the profit uncertainty is more detrimental to firms with higher liability ratios, and that the estimated size of the uncertainty effect (in an absolute value) increases monotonically. Assuming that the liability ratio reflects marginal costs for external fundings, this suggests that the profit uncertainty is

expected to contribute more strongly to a firm's investment for a firm with higher external funding costs.

Industry

Lastly, we consider the firms' sub-industry classification within the manufacturing industry, which is a frequently used grouping criterion in previous studies. Assuming that firms within the same sub-industry share common characteristics such as the price elasticity of demand and production technology, we divide the sample into 20 groups based on the first 2-digit SIC code, following [57]. For instance, the SIC code for Texaco Inc. is 2911, with the first two digits (29) indicating that its core business is related to "petroleum and coal products.". The first three columns in Panel 3 of Table 3.3 provide the 2-digit SIC code, the description of the core businesses and the number of firms within each group.

Sub-industry level estimates of γ_i's are shown in the last column in Panel 3 of Table 3.3 as well as in Figure 3.7. They are again substantially different from each other. For instance, the point estimate of "furniture and fixture (SIC2=25)" is -0.4258, indicating that investment of this industry is expected to decline facing high profit uncertainty. On the contrary, firms in the "electronic and other electric equipment (SIC2=36)" industry rather increase capital expenditure when uncertainty is high ($\gamma = 0.0778$).

Regarding the sub-industry classification, Leahy and Whited [79] compare the labor-to-capital ratios of sub-industries. Theoretically, a higher or more volatile labor-to-capital ratio means a lower capital intensity and thus a higher substitutability of capital by labor. From this perspective, investment activities of firms facing such flexible production technologies would be less affected by profit uncertainty.[13] However, Leahy and Whited found evidence contradicting the theoretical implication that firms with a higher (and/or more volatile) labor-to-capital ratio

---

[13]The assumption that a higher labor-to-capital ratio (or higher volatility of the ratio) can be attributed to an easier substitutability of capital holds when a firm faces a convex return, as noted in [1].

would reduce capital investment more, using a two-group (high/low substitutability) approach.

Following Leahy and Whited [79], we calculate the labor-to-capital ratio of sub-industries in the year 2000, as well as the historical standard deviation of yearly changes in Panel 3 of Table 3.3.[14] Further, Figures 3.8 and 3.9 show the coefficients plotted against the labor-to-capital ratios and their historical standard deviations, respectively. Both figures show overall negative correlations between γ_i's and the labor-to-capital ratios (or standard deviations), in line with the findings of Leahy and Whited. However, our approach provides a more detailed picture of the relation, ase it examines 20 sub-industry groups.

Therefore, we infer that the labor-to-capital ratio or its historical variability is not a direct indicator of the labor-capital substitutability and, further, the convexity of the production function. Related to this point, it is possible for an industry to be capital-intensive, although the substitutability between labor and capital is still high. For example, a firm within the industry with a high labor-to-capital ratio, such as "transportation equipment (SIC2=37)," would still have relatively high fixed costs of capital, making it difficult to cope with high profit uncertainty.

3.4.4 Low-Frequency Movements of Uncertainty

Our analysis so far was based on a one-quarter-ahead profit uncertainty measure. One remaining concern is whether the result is driven by the frequent volatile fluctuations of our estimated profit uncertainty. Therefore, we investigate whether the result still holds when low-frequency dynamics in the aggregate profit

---

[14]The historical labor and capital productivity indexes are obtained from the Bureau of Labor Statistics, "Superseded historical SIC measures for manufacturing sectors and 2-digit SIC manufacturing industries, 1949-2001." The labor-to-capital ratio in the year 2000 is used as a representative level, and the historical standard deviation, calculated from yearly log changes, is used as a proxy for a variation for each sub-industry.

uncertainty measure is used.[15]

We extract low-frequency dynamics of our profit uncertainty measure $(\bar{\sigma}^2_{t+1|t})$ by taking moving averages with five-quarter rolling windows:

$$\bar{\sigma}^2_{t+1|t} \equiv \frac{\sigma^2_{t-1|t-2} + \sigma^2_{t|t-1} + \sigma^2_{t+1|t} + \sigma^2_{t+2|t+1} + \sigma^2_{t+3|t+2}}{5}. \qquad (3.10)$$

In Figure 3.10, $\bar{\sigma}^2_{t+1|t}$ shows much smoother dynamics, with most short-run fluctuations having disappeared. Table **??** reports estimation results of our baseline model under the alternative measure. Here we also include the result using recursive four-quarter-ahead forecasts of profit uncertainty for comparison.[16]

We find that baseline result is robust under both alternative measures of the profit uncertainty. If anything, the size of coefficient estimates increases in absolute terms. Related to this, Alquist et al. [4] and Kilian and Vigfusson [75] point out that investment, particularly the one incurring large-scale expenditures, is affected much more by uncertainty at a longer forecast horizons rather than its short-lived quarterly fluctuations. Hence, assuming that the alternative measures of the low-frequency movement in uncertainty reflect changes in longer horizon than a quarter, the more destructive result we find is in line with this argument.[17]

3.5 Conclusion

This paper investigates how aggregate profit uncertainty affects manufacturing firms' investment activities, accounting for heterogeneous investment ad-

---

[15]We appreciate John Campbell's suggestion regarding the importance of the low-frequency movement in future profit uncertainty in firm investment activity.

[16]Conditional on information at time t, recursively m-period-ahead uncertainty forecasts $(\sigma^2_{t+m|t})$ are generated, as $\sigma^2_{t+m|t} = E\left[\left(\sum_{i=1}^{J} w^j_{t+m-1} \cdot u^j_{t+m}\right)^2 |F_t\right] = \alpha_0 + \alpha_1 \cdot \sigma^2_{t+m-1|t}$. In particular, we use a four-quarter-ahead forecast of the profit uncertainty, $\sigma^2_{t+4|t}$.

[17]The more negative coefficient estimates is also related to the lower long-run average of those measures. In particular, the four-quarter-ahead uncertainty series generated recursively has a level and magnitude smaller than one-quarter-ahead uncertainty, due to the fact that in an ARCH model the volatility converges to a long-run mean, but the overall dynamics remain the same as before.

justments across firms. A large body of literature have examined the effect of uncertainty, but their findings so far differ greatly. Therefore, by highlighting the firm-level heterogeneity, this paper attempts to provide a consolidating ground in understanding the various channels of uncertainty transmission explored in the previous literature.

We propose an investment forecasting model that coherently assesses the heterogeneous adjustments of firms' business investment in responding to changes in profit uncertainty: the model includes profit uncertainty interacted with firm-specific slope coefficients. As a consequence, the model can flexibly investigate potential heterogeneity at the firm level, without having to divide samples in a certain way a priori, contrary to previous studies. Moreover, framework easily detects any non-linear or non-monotonic relationship between a certain firm characteristic and the size of uncertainty effects.

Another methodological contribution of this paper is to introduce a Panel-ARCH model for aggregate uncertainty. We model aggregate profit uncertainty to be the conditional standard deviation of the common factor that simultaneously drives unpredicted variations in a large panel of the industry-level sales growth. We further postulate that factor loadings and idiosyncratic volatilities are inversely proportional to an industry's sales share, which enables the derivation of closed-form log-likelihood function as well as the aggregate volatility process. Thus, the Panel-ARCH model makes it simple to utilize information in a large cross-section for the volatility estimation. The proposed model can be further applied for an estimation of another aggregate volatility series; in particular, when the series of interest is only available at a low frequency with a short history, the estimation of its volatility may not be very simple. In this case, if its subcomponents are still attainable, one can apply the Panel-ARCH model and make use of cross-sectional variations for the estimation of aggregate uncertainty through a factor-structure.

Using the Compustat data from 1989Q1 to 2012Q4, we find that aggregate profit uncertainty affects firms' investment activity slightly negatively on average,

consistent with the views in [9] and [12], among others. However, we find that firms differ substantially in the way they adjust capital investment in response to changes in aggregate profit uncertainty. We further investigate several firm characteristics attributable to the heterogeneity. In particular, we observe an inverse U-shape in the magnitude of uncertainty effects in relation to firm size which has not been documented in previous literature such as Ghosal and Loungani [53]. Therefore, it will be interesting for future studies to examine the non-linear relation found in our paper further, linking it to the model shown in Bolton et al. [14], which demonstrates that a firm's internal funds and investment timing has a highly non-linear and non-monotonic relationship.

3.6   Acknowledgements

Chapter 3, "Heterogeneity in the Dynamic Effects of Uncertainty on Investment", is coauthored with Soojin Jo, and in full, has been submitted for publication of material. I thank Soojin for the permission to use this chapter in my dissertation.
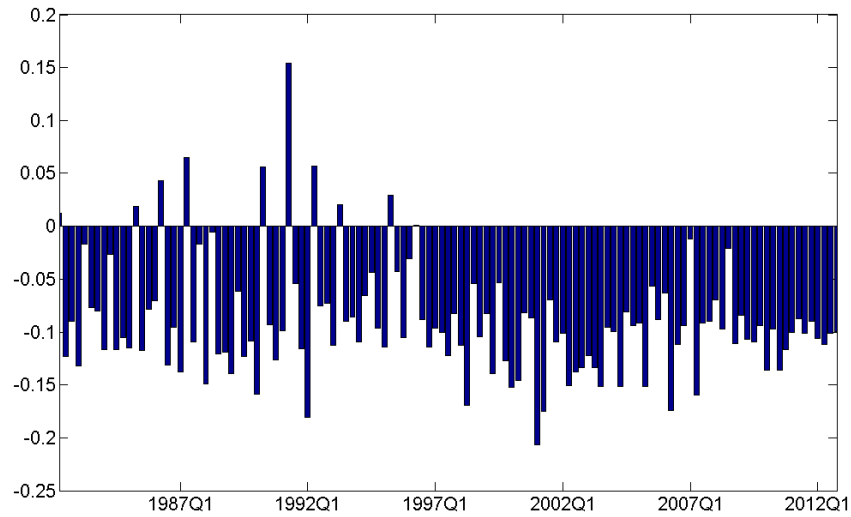
Figures and Tables

Figure 3.1: Correlation between Size of Forecasting Errors and Sales Shares

Note: This figure shows the correlation coefficients between the absolute values of fore-casting errors from the sales growth forecasting equation (3) and the previous quarter's sales shares in each quarter t. Each period, we have forecasting errors of 67 industries $(u_{j,t+1})$ and their sales series from the previous quarter $(w_{j,t})$, with which we compute correlation coefficients. This implies that an industry with a smaller sales share tens to have a larger unpredicted variation, supporting our specification in (6).
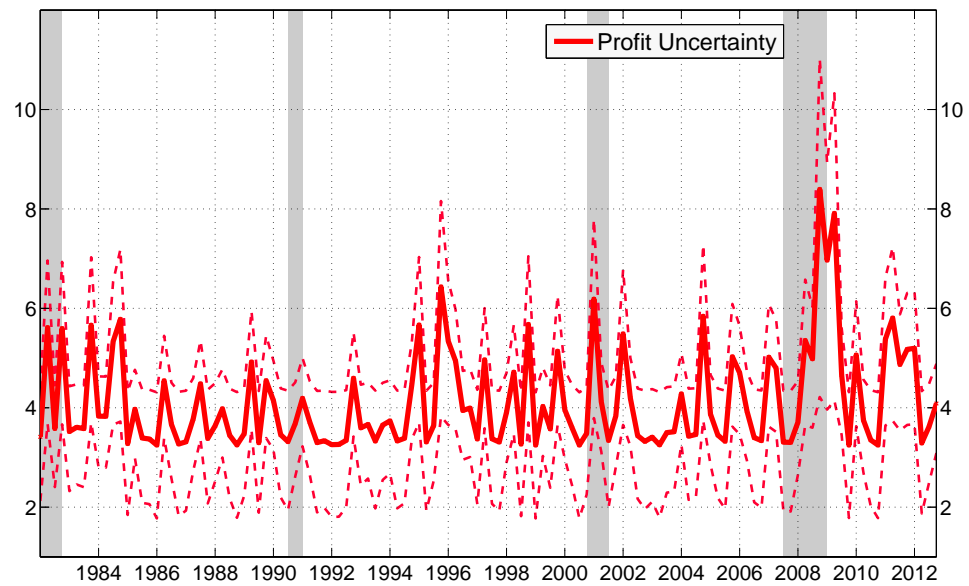
Figure 3.2: Aggregate Profit Uncertainty Estimated from a Panel-ARCH(1)

Note: Plotted above is the estimated aggregate profit uncertainty series from the Panel-ARCH(1) model from 1982Q1 to 2012Q4. The dotted lines represent lower and upper bounds for the 95% confidence band, constructed by simulating the Panel-ARCH(1) model 1,000 times using MLE estimates and the Hessian matrix. Shaded are the NBER recession dates.
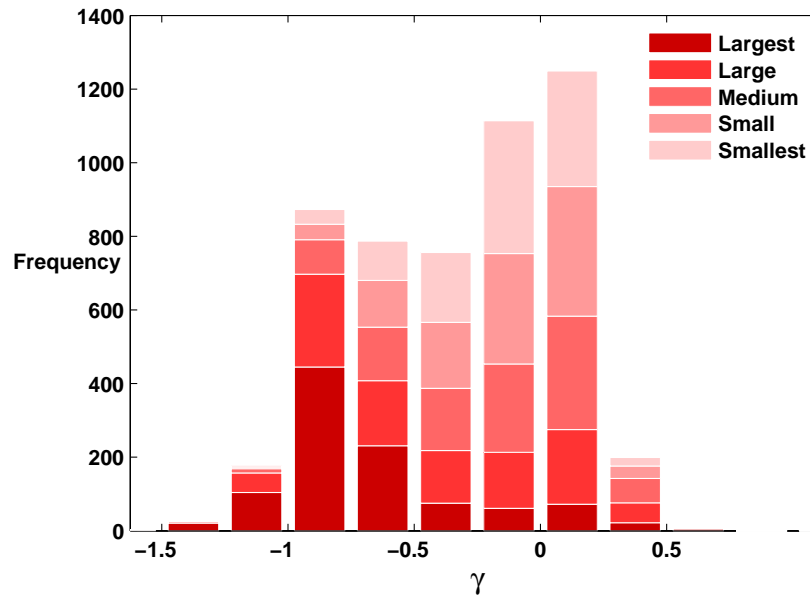
Figure 3.3: Responses to Uncertainty Grouped by Firm Size

Note: This is a histogram of the firm-specific slope coefficient estimates ($\hat{\gamma}_i$s) across $5,197$ firms, from the model with complete firm-level heterogeneity: each firm is allowed to respond differently to aggregate profit uncertainty. The bars are color-coded to represent the size quintile to which a firm belongs.
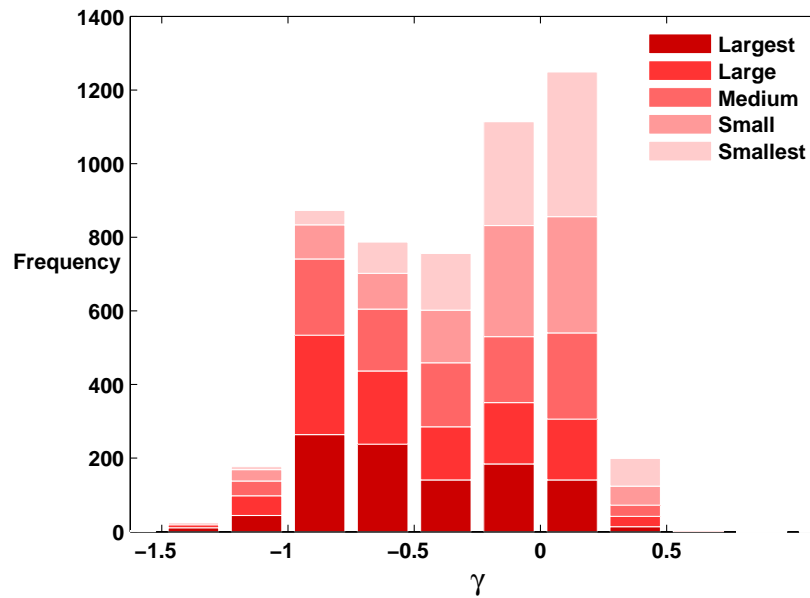
Figure 3.4: Responses to Uncertainty Grouped by Liability Ratio

Note: This is a histogram of the firm-specific slope coefficient estimates ($\hat{\gamma}_i$'s) across 5,197 firms, from the model with complete firm-level heterogeneity: each firm is allowed to respond differently to aggregate profit uncertainty. It is the same histogram as Figure 3.3, but the bars are color-coded to represent the quintiles of the total-liabilities-to-asset ratio to which a firm belongs.
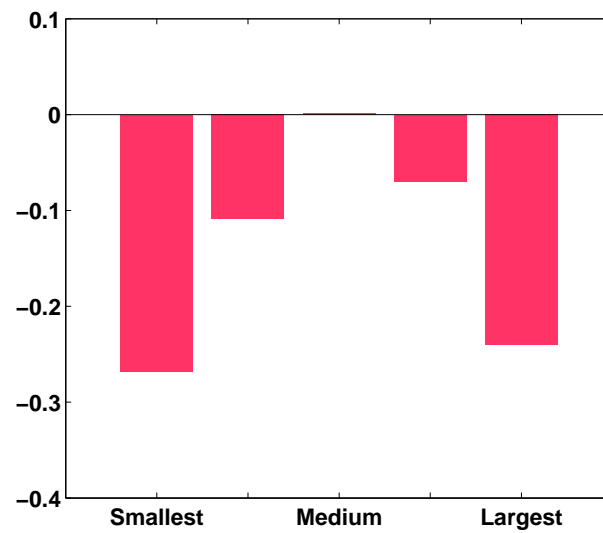
Figure 3.5: By Firm Size: Responses to Uncertainty

Note: This figure plots the estimates of group-specific slope coefficients. Firms are first divided into quintiles according to size; then firms of the same group are assumed to share the same slope, which differs from that of other groups.
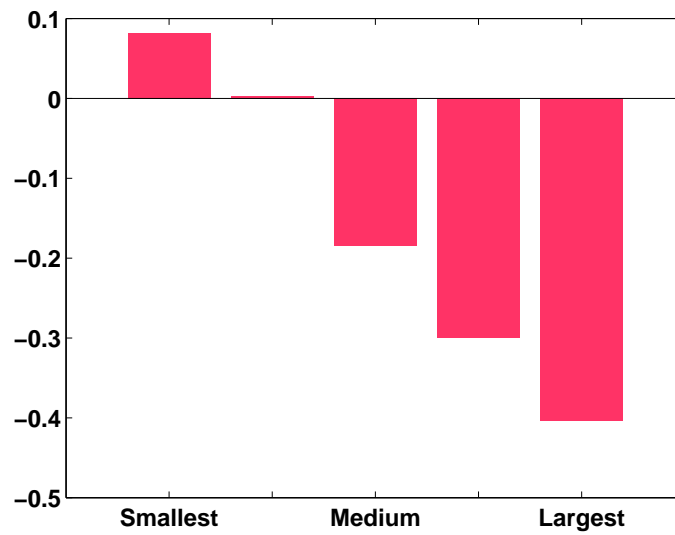
Figure 3.6: By Liability Ratio: Responses to Uncertainty

Note: This figure plots the estimates of group-specific slope coefficients. Firms are first divided into quintiles based on their liabilities-to-asset ratio; then firms of the same group are assumed to share the same slope, which differs from that of other groups.
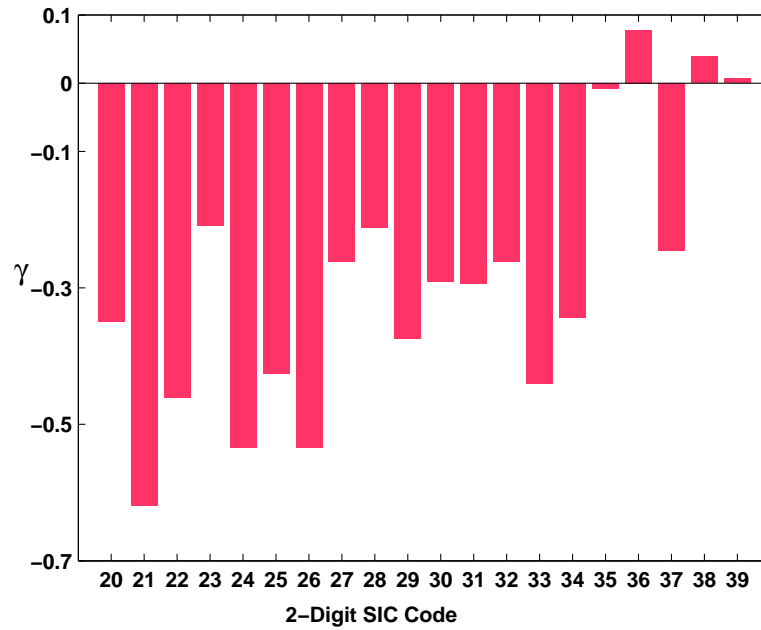
Figure 3.7: By Sub-Industry: Responses to Uncertainty

Note: This figure plots the estimates of group-specific slope coefficients. Firms are first divided into groups based on their 2-digit SIC code; firms of the same group are assumed to share the same slope, which differs from that of other groups.

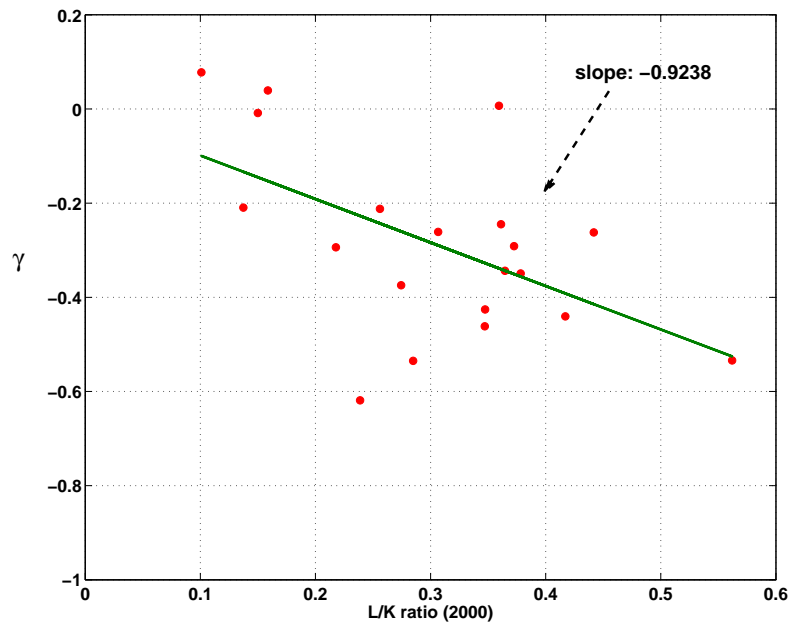Figure 3.8: Responses to Uncertainty by Labor-to-Capital Ratios

Note: This is a scatter plot of 2-digit sub-industry-specific slope coefficient estimates versus labor-to-capital ratios of sub-industries. The labor-to-capital ratios are measured in the year 2000. For demonstration purposes, we plot the fitted line from regressing sub-industry-specific slope coefficient estimates on a constant and labor-to-capital ratios.
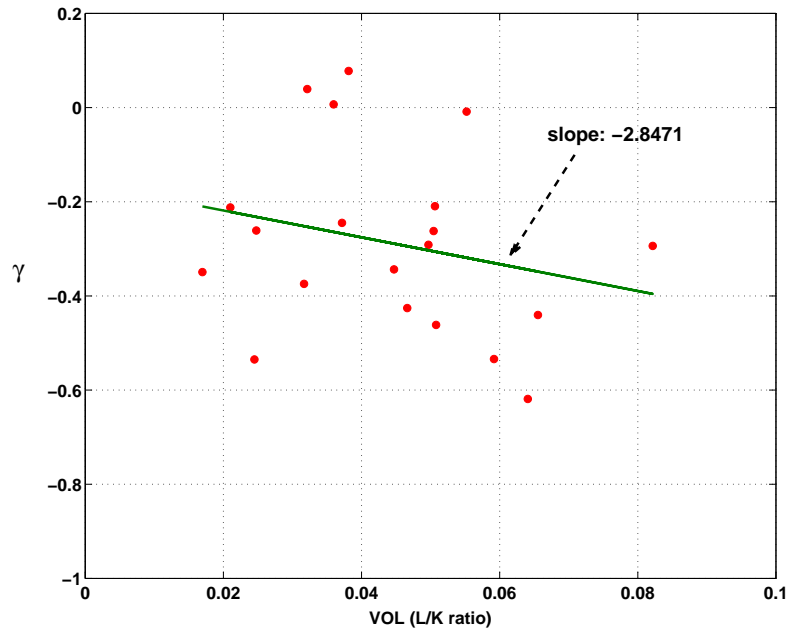
Figure 3.9: Responses to Uncertainty by Standard Deviations of Labor-to-Capital Ratios

Note: This is a scatter plot of 2-digit sub-industry-specific slope coefficient estimates versus standard deviations of labor-to-capital ratios of sub-industries. The standard deviations of the labor-to-capital ratios are calculated from changes in the logarithm of the ratio over time. For demonstration purposes, we plot the fitted line from regressing sub-industry-specific slope coefficient estimates on a constant and standard deviations of labor-to-capital ratios of sub-industries.

Figure 3.10: Five-Quarter Moving Average of Profit Uncertainty

Note: Plotted above is the low-frequency dynamics of profit uncertainty calculated as moving averages of a five-quarter rolling window (in blue), along with the baseline profit uncertainty series (in red).

Table 3.1: Descriptive Statistics

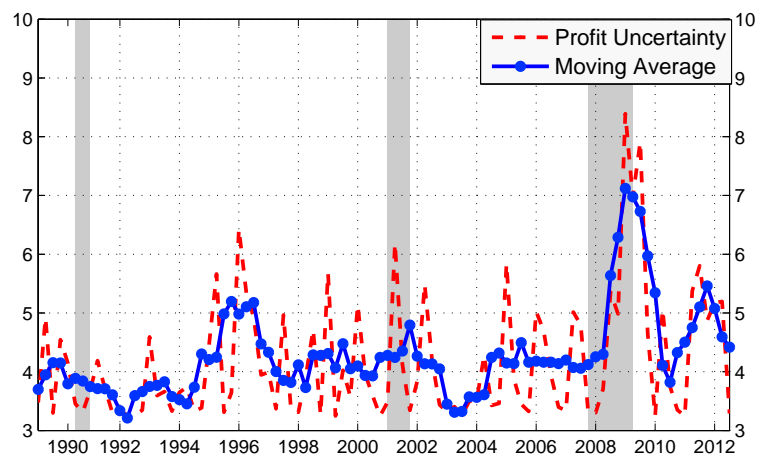| Variable	Description	Mean	S.D.	Skew	Kurt	Min	Percentile			
							25%	50%	75%	Max
investment (y_{t+1})	investment-to-capital-stock ratio $=i_{t+1}/k_t$	0.06	0.07	2.35	7.18	0.00	0.02	0.04	0.08	0.58
sales	sales-to-capital-stock ratio $=s_t/k_{t-1}$	1.90	2.20	3.27	15.93	0.00	0.66	1.22	2.25	25.73
cash	cash-to-capital-stock ratio $=ce_t/k_{t-1}$	2.72	7.51	7.14	77.95	0.00	0.08	0.40	1.92	177.00
size	log total book-value asset $=\log a_t$	5.21	2.29	0.20	−0.15	−4.17	3.60	5.10	6.73	13.10
liability ratio	total-liabilities-to-total-assets ratio $=l_t/a_t$	0.50	0.34	2.63	16.87	0.00	0.27	0.47	0.65	6.53

Note: This table provides the summary statistics of key variables. The sample period runs from 1989Q1 to 2012Q4 (96 quarters) and the sample contains 5,197 firms with 219,538 firm-quarter observations. We obtain firm-quarter capital expenditure (i), net property plant and equipment (k), sales/turnover (s), cash and cash equivalent (ce), total book-valued assets (a), and total book-valued liabilities (l) from the Compustat, and convert them into the above ratios.

Table 3.2: Parameter Estimates

	(1) No Heterogeneity		(2) Firm-Level		(3) Size		(4) Liability Ratio		(5) Industry	
	MLE	(s.e.)	MLE	(s.e.)	MLE	(s.e.)	MLE	(s.e.)	MLE	(s.e.)
Uncertainty (γ)	-0.1387^{***}	(0.0134)								
Constant (β_0)	4.5637^{***}	(0.0720)	4.2615^{***}	(0.0390)	4.4847^{***}	(0.0914)	4.3327^{***}	(0.0717)	4.3752^{***}	(0.0720)
Sales (β_1)	0.0061^{***}	(0.0001)	0.0054^{***}	(0.0001)	0.0061^{***}	(0.0001)	0.0061^{***}	(0.0001)	0.0056^{***}	(0.0001)
Cash (β_2)	0.0011^{***}	(0.0000)	0.0008^{***}	(0.0000)	0.0011^{***}	(0.0000)	0.0009^{***}	(0.0000)	0.0011^{***}	(0.0000)
Size (β_3)	0.0057	(0.0060)	0.3073^{***}	(0.0057)	0.0240	(0.0126)	0.0709^{***}	(0.0061)	0.0692^{***}	(0.0062)
Lag (β_4)	0.0612^{***}	(0.0007)	0.0535^{***}	(0.0006)	0.0613^{***}	(0.0007)	0.0604^{***}	(0.0007)	0.0596^{***}	(0.0007)
GDP (δ)	0.5971^{***}	(0.0220)	0.4694^{***}	(0.0199)	0.5860^{***}	(0.0220)	0.6113^{***}	(0.0219)	0.6129^{***}	(0.0218)
ζ	6.1275^{***}	(0.0094)	5.8395^{***}	(0.0089)	6.1124^{***}	(0.0093)	6.0846^{***}	(0.0093)	6.0782^{***}	(0.0093)
Likelihood	$-496,176.86$		$-485,848.38$		$-495,648.22$		$-494,670.31$		$-494,444.09$	

Note: Table 3.2 reports coefficient estimates and their asymptotic standard errors of the baseline model without heterogeneity (column (1)), the model with complete firm-level heterogeneity (column (2)), and the model with group-level heterogeneity (columns (3)-(5)). The reported asymptotic standard errors are estimated by numerically approximating the second derivatives of the Hessian matrix at the MLE. The asterisks, */**/***, represent the statistical significance at 90%, 95% and 99%, respectively.

Table 3.3: Group Characteristics and Slope Coefficient Estimates

Panel 1: Size group description				γ	
	Description	No.	Size	MLE	(s.e.)
1	Smallest	1039	1.90	-0.2681^{***}	(0.0177)
2		1040	3.58	-0.1087^{***}	(0.0156)
3		1039	4.68	0.0011	(0.0149)
4		1040	5.88	-0.0699^{***}	(0.0149)
5	Largest	1039	8.03	-0.2402^{***}	(0.0166)
Panel 2: Liability ratio group description				γ	
	Description	No.	Liability Ratio	MLE	(s.e.)
1	Smallest	1039	0.20	0.0817^{***}	(0.0147)
2		1040	0.35	0.0025	(0.0146)
3		1039	0.49	-0.1843^{***}	(0.0145)
4		1040	0.62	-0.2994^{***}	(0.0148)
5	Largest	1039	0.94	-0.4037^{***}	(0.0151)

Note: Table 3.3 describes sample-split criteria, their main characteristics (Description) as well as the number of firms in each group (No.). Panels 1 and 2 report historical averages of book-valued assets (in logarithm), and total-liabilities-to-total-assets ratio, respectively. Panel 3 reports the labor-to-capital ratio (L/K) in year 2000, as well as its historical volatility (Vol) from 1980 to 2001. The last column reports estimated group-specific coefficients (γ_i's) and their asymptotic standard errors of each group. The asterisks, */**/***, represent the statistical significance at 90%, 95% and 99%, respectively.

Table 3.3 : Group Characteristics and Slope Coefficient Estimates, continued

	Panel 3: SIC2 group description				γ	
	Description	No.	L/K	Vol	MLE	(s.e.)
20	Food and kindred	264	0.38	0.02	-0.3494^{***}	(0.0186)
21	Tobacco	10	0.24	0.06	-0.6190^{***}	(0.0677)
22	Textile mill	72	0.35	0.05	-0.4616^{***}	(0.0315)
23	Apparel and other textile	99	0.14	0.05	-0.2095^{***}	(0.0256)
24	Lumber and wood	70	0.56	0.06	-0.5341^{***}	(0.0282)
25	Furniture and fixtures	60	0.35	0.05	-0.4258^{***}	(0.0292)
26	Paper	126	0.28	0.02	-0.5349^{***}	(0.0234)
27	Printing and publishing	142	0.31	0.02	-0.2611^{***}	(0.0231)
28	Chemicals	1020	0.26	0.02	-0.2122^{***}	(0.0148)
29	Petroleum and coal	82	0.27	0.03	-0.3745^{***}	(0.0274)
30	Rubber and plastic	147	0.37	0.05	-0.2914^{***}	(0.0229)
31	Leather	26	0.22	0.08	-0.2939^{***}	(0.0397)
32	Stone, clay and glass	64	0.44	0.05	-0.2623^{***}	(0.0300)
33	Primary metal industries	190	0.42	0.07	-0.4406^{***}	(0.0207)
34	Fabricated metal	172	0.36	0.04	-0.3437^{***}	(0.0214)
35	Industrial machinery	682	0.15	0.06	-0.0085	(0.0157)
36	Electronic and other electric	889	0.10	0.04	0.0778^{***}	(0.0148)
37	Transportation equipment	240	0.36	0.04	-0.2449^{***}	(0.0187)
38	Instruments	720	0.16	0.03	0.0393^{***}	(0.0155)
39	Miscellaneous	122	0.36	0.04	0.0069	(0.0248)

Table 3.4: Baseline Model Estimation Results with Low-Frequency Movement of Uncertainty

	(1) moving average ($\bar{\sigma}^2_{t+1\|t}$)		(2) 1-year ahead ($\sigma^2_{t+4\|t}$)	
	MLE	(s.e.)	MLE	(s.e.)
Uncertainty (γ)	-0.7178***	(0.0288)	-12.1756***	(1.0985)
Constant (β_0)	6.8848***	(0.1246)	48.5089***	(4.0203)
Sales (β_1)	0.0061***	(0.0001)	0.0061***	(0.0001)
Cash (β_2)	0.0011***	(0.0000)	0.0011***	(0.0000)
Size (β_3)	0.0039	(0.0060)	0.0056	(0.0060)
Lag (β_4)	0.0609***	(0.0007)	0.0612***	(0.0007)
GDP (δ)	0.6135***	(0.0210)	0.5762***	(0.0225)
ζ	6.1202***	(0.0094)	6.1273***	(0.0094)
Likelihood	-495,920.17		-496,168.70	

Note: Table 3.4 reports estimation results of our baseline model with low-frequency variation of profit uncertainty. Column (1) is the result using moving averages of a five-quarter rolling window, as in Equation (3.10). For comparison, we also present the results using a four-quarter-ahead profit uncertainty forecast generated recursively. The statistical significance of the MLE estimates is indicated by asterisks, */**/***, representing the statistical significance at 90%, 95% and 99%, respectively.

Appendix 3.A: Panel-ARCH for Aggregate Profit Uncertainty

3.A.1 The log likelihood function of Panel-ARCH

Let $\Theta \equiv [a_0, a_1, \cdots, a_q, \tau]$ be the collection of parameters to estimate. With u_{t+1} denoting a $J \times 1$ vector of forecasting errors, the log-likelihood function jointly modeling forecasting errors of J industries for $t = 1, ..., T$ is

$$L\left(\Theta\right) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log |\Omega_{t+1}| - \frac{1}{2}\sum_{t=1}^{T} u'_{t+1}\Omega_{t+1}^{-1}u_{t+1}, \qquad (3.11)$$

where $\Omega_{t+1} = E\left[u_{t+1}u'_{t+1}|\mathcal{F}_t\right]$ is a $J \times J$ variance-covariance matrix of forecasting errors.

The evaluation of the above log-likelihood function with a large J is generally burdensome, as it requires the determinant calculation and inversion of $J \times J$ matrix Ω_{t+1} in each point t. However, since Ω_{t+1} is symmetric with all elements known under condition (3.6) for all t, its determinants and inverses are analytically derived. More specifically, the matrix Ω_{t+1} has a form as

$$\Omega_{t+1} = \begin{bmatrix} \frac{\sigma_{t+1}^2+\tau^2}{J^2 \cdot w_{1,t}^2} & \frac{\sigma_{t+1}^2}{J^2 w_{1,t}w_{2,t}} & \cdots & \frac{\sigma_{t+1}^2}{J^2 w_{1,t}w_{J,t}} \\ \frac{\sigma_{t+1}^2}{J^2 w_{2,t}w_{1,t}} & \frac{\sigma_{t+1}^2+\tau^2}{J^2 \cdot w_{2,t}^2} & \cdots & \frac{\sigma_{t+1}^2}{J^2 w_{2,t}w_{J,t}} \\ \vdots & \vdots & & \vdots \\ \frac{\sigma_{t+1}^2}{J^2 w_{J,t}w_{1,t}} & \frac{\sigma_{t+1}^2}{J^2 w_{J,t}w_{2,t}} & \cdots & \frac{\sigma_{t+1}^2+\tau^2}{J^2 \cdot w_{J,t}^2} \end{bmatrix}.$$

Thus, the determinant of Ω_{t+1} can be analytically calculated as

$$|\Omega_{t+1}| = \tau^{2(J-1)} \cdot \left(\tau^2 + J\sigma_{t+1}^2\right) \cdot \left(\prod_{j=1}^{J} \frac{1}{J^2 w_{j,t}^2}\right).$$

In addition, its inverse matrix is summarized by

$$
\Omega_{t+1}^{-1}(i,j) = \begin{cases} \frac{J^2 w_{i,t}^2 \cdot \left(\tau^2 + (J-1)\cdot\sigma_{t+1}^2\right)}{\tau^2 \cdot \left(\tau^2 + J\cdot\sigma_{t+1}^2\right)} & \text{for } j = i \\ -\frac{J^2 w_{i,t} w_{j,t} \cdot \sigma_{t+1}^2}{\tau^2 \cdot \left(\tau^2 + J\cdot\sigma_{t+1}^2\right)} & \text{for } j \neq i \end{cases}
$$

where $\Omega_{t+1}^{-1}(i,j)$ an $(i,j)^{th}$ element in Ω_{t+1}^{-1}.

As a result, the closed-form log-likelihood function can be derived by plugging the above analytic formula into the general log-likelihood function (3.11) as follows:

$$
\begin{aligned}
L(\Theta) = & -\frac{T}{2}\log 2\pi - T(J-1)\log(\tau) + 2T\log(J) \\
& -\frac{1}{2}\sum_{t=1}^{T}\log\left(\tau^2 + J\sigma_{t+1}^2\right) + \sum_{t=1}^{T}\sum_{j=1}^{J}\log w_{j,t} \\
& -\frac{J^2}{2\tau^2}\sum_{t=1}^{T}\left\{\sum_{j=1}^{J}(w_{j,t}u_{j,t+1})^2 - \frac{\sigma_{t+1}^2}{\left(\tau^2 + J\cdot\sigma_{t+1}^2\right)}\cdot\left(\sum_{i=1}^{J}w_{j,t}u_{j,t+1}\right)^2\right\}.
\end{aligned} \tag{3.12}
$$

Evaluating the above function hence becomes as simple as that of a univariate series. Therefore, the parsimonious specifications of the Panel-ARCH provides tractability, overcoming empirical difficulties due to a large J.

3.A.2 Aggregate sales revenue index and its growth rate

Figure 3.A.1 plots the quarterly sales revenue index calculated as (3.1) after setting the index level in 1981Q1=100; the gross sales revenue (S_t) increases over time with strong seasonality. After controlling for the observed seasonality in S_t through seasonal dummies, we calculate the quarterly sales growth rate as plotted in Figure 2.A.2. The sales growth is relatively low during recessions, but overall very volatile, with volatility clustered similar to the behavior of the stock return.
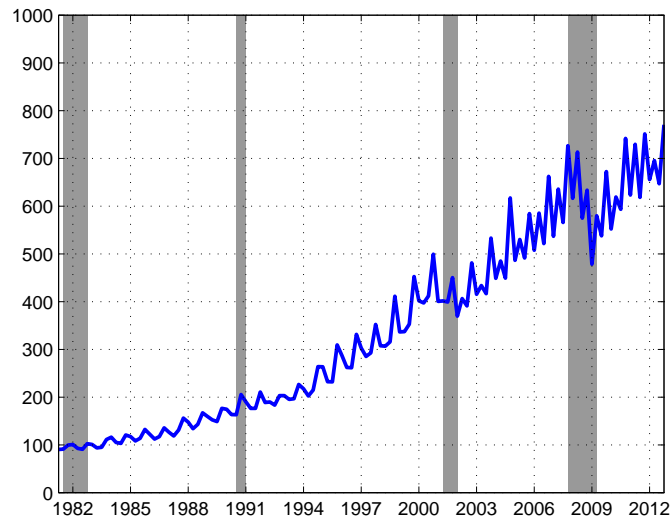
Figure 3.A.1: Aggregate Sales Index

Note: This figure plots the aggregate sales revenue index from 1981Q1 to 2012Q4. The series is normalized by setting the initial level equal to 100.
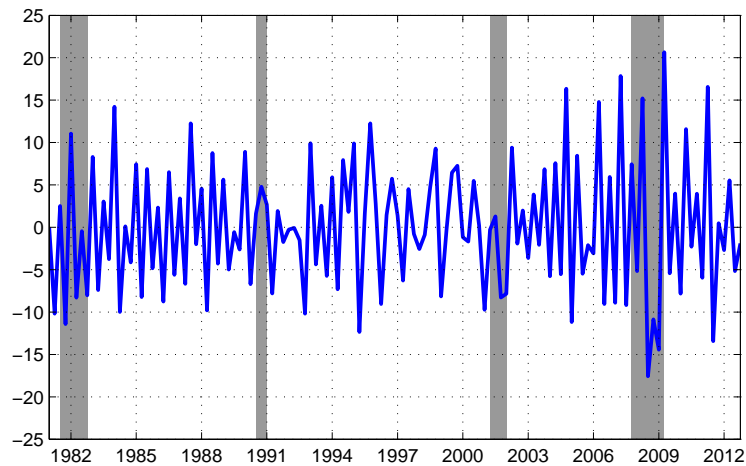

Figure 3.A.2: Quarterly Sales Growth Rate

Note: This figure plots the de-seasonalized quarterly growth rate of the aggregate sales index. De-seasonalization is done by regressing the growth rate on seasonal dummies.

3.A.3 Panel-ARCH model specification

When estimating aggregate profit uncertainty from a panel of the industry-level sales growth forecasting errors from 1982Q1 to 2012Q4 (124 quarters), we consider three different specifications of the Panel-ARCH model: Panel-ARCH(1), Panel-ARCH(2) and Panel-GARCH(1,1) as an extension to our general framework. For model parameter estimation, we condition on the first 4 quarter observations.

Table 3.A.1 reports maximum likelihood estimates, their asymptotic standard errors, and maximized log-likelihoods under the three specifications. When comparing maximized log-likelihoods, we find that the Panel-ARCH(1) is parsimonious yet sufficient to describe the dynamics of profit uncertainty: the p-values for log-likelihood ratio tests are 0.47 for the Panel-ARCH(2) and 0.43 for the Panel-GARCH(1,1). Hence, we use the estimated profit uncertainty from the Panel-ARCH(1) model in the paper.

Table 3.A.1: Order Determination of Panel-ARCH Model

	(1) Panel-ARCH(1)		(2) Panel-ARCH(2)		(3) Panel-GARCH(1,1)	
	MLE	(s.e.)	MLE	(s.e.)	MLE	(s.e.)
α_0	10.5412***	(4.7540)	9.4208**	(4.9186)	4.7165	(8.9977)
α_1	0.2085**	(0.1210)	0.1816*	(0.1297)	0.1658	(0.1312)
α_2			0.0506	(0.0819)		
β_1					0.3981	(0.6628)
τ	33.6701***	(0.2632)	33.6701***	(0.2632)	33.6701***	(0.2632)
Likelihood	-43,284.60		-43,284.34	(0.47)	-43,284.29	(0.43)

Note: We report MLE estimates and maximized log-likelihoods under various order specifications. Column (3) shows the results under the Panel-GARCH(1,1) model, where β_1 is the coefficient for lagged uncertainty. The asymptotic standard errors are obtained by approximating the second derivative of the Hessian matrix. Statistical significance of MLE estimates is indicated by using asterisks */**/***, representing the statistical significance at 90%, 95% and 99%, respectively. The numbers in parentheses next to log-likelihoods are the p-values for a likelihood ratio test of one model against the Panel-ARCH(1) model, where the test statistic is twice the difference in log-likelihoods and is distributed chi-squared with degrees of freedom equal to the difference in the number of parameters.

Appendix 3.B: Data for Investment Forecasting Model

The manufacturing firm panel are taken from the combined quarterly Compustat North America industrial file, from which we obtain capital expenditure (investment, i_t), net property plant and equipment (capital, k_t), total book-valued assets (size, a_t), sales/turnover (sales, s_t), cash and cash equivalents (cash, ce_t), and total liabilities (liability, l_t). All variables except the capital expenditure are reported on a quarterly basis; for example, capital is the book value of capital stocks for the reported quarter, and sales refers to the gross revenue from goods sold during the quarter. The capital expenditure, however, is reported as a year-to-date item, which we transform to be a quarterly value by subtracting the previous period's (year-to-date) amount from that of the current period.

Given the unbalanced panel data spanning from 1989Q4 to 2012Q4, we first delete observations with missing variables, non-positive capital and size, negative sales and/or cash, and with liabilities likely to indicate highly distressed firms. We also delete observations with capital expenditures greater than 15% of book value assets, following Leahy and Whited [79], and observations either below 2.5% or above 97.5% in each quarter for selected variables: i_{t+1}/k_t, s_t/k_{t-1}, ce_t/k_{t-1} and l_t/a_t. Lastly, we remove 1,981 firms with less than 12 quarterly observations during the sample periods, corresponding to 11,384 firm-quarter observations.

Due to different fiscal-year conventions across firms, we create quarterly observations based on their reporting dates. To normalize units, we divide the capital expenditure (i_{t+1}), sales (s_t), and cash (ce_t) by the beginning-period capital stocks. Lastly, we adjust the total book value assets (a_t) in 2005Q3 US dollars.

References

[1] Abel, A. B., 1983: Optimal investment under uncertainty. *American Economic Review*, **73**(1), 228–233.

[2] Acharya, V. V., Almeida, H., and Campello, M., 2007: Is cash negative debt? a hedging perspective on corporate financial policies. *Journal of Financial Intermediation*, **16**(4), 515–554.

[3] Acharya, V. V., Lochstoer, L. A., and Ramadorai, T., 2013: Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics*, **109**(2), 441–465.

[4] Alquist, R., Kilian, L., and Vigfusson, R., forthcoming: Forecasting the price of oil. In *Handbook of Economic Forecasting, 2, edited by Graham Elliott and Allan Timmermann*. Amsterdam: North-Holland.

[5] Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X., 2006: The cross-section of volatility and expected returns. *Journal of Finance*, **61**(1), 259–299.

[6] Asker, J., Farre-Mensa, J., and Ljungqvist, A., 2011: Comparing the investment behavior of public and private firms. *NBER Working Paper No. 17394*.

[7] Aulerich, N., Irwin, S., and Garcia, P., 2010: The price impact of index funds in commodity futures markets: evidence from the cftc's daily large trader reporting system. *University of Illinois at Urbana-Champaign*. Working paper.

[8] Baum, C. F., Caglayan, M., and Talavera, O., 2008: Uncertainty determinants of firm investment. *Economics Letters*, **98**(3), 282–287.

[9] Bernanke, B. S., 1983: Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics*, **98**(1), 85–106.

[10] Bernanke, B. S., Gertler, M., and Gilchrist, S., 1999: The financial accelerator in a quantitative business cycle framework. *Handbook of Macroeconomics*, **1**, 1341–1393.

[11] Bloom, N., 2009: The impact of uncertainty shocks. *Econometrica*, **77**(3), 623–685.

[12] Bloom, N., Bond, S., and Van Reenen, J., 2007: Uncertainty and investment dynamics. *Review of Economic Studies*, **74**(2), 391–415.

[13] Bollerslev, T., 1986: Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, **31**(3), 307–327.

[14] Bolton, P., Wang, N., and Yang, J., 2014: Investment under uncertainty and the value of real and financial flexibility.

[15] Boyle, G. W., and Guthrie, G. A., 2003: Investment, uncertainty, and liquidity. *Journal of Finance*, **58**(5), 2143–2166.

[16] Brailsford, T. J., and Faff, R. W., 1996: An evaluation of volatility forecasting techniques. *Journal of Banking and Finance*, **20**(3), 419–438.

[17] Brennan, M., 1958: The supply of storage. *American Economic Review*, **48**(1), 50–72.

[18] Brownlees, C., Engle, R., and Kelly, B., 2012: A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, **14**(2), 3.

[19] Bryant, H., Bessler, D., and Haigh, M., 2006: Causality in futures markets. *Journal of Futures Markets*, **26**(11), 1039–1057.

[20] Büyüksahin, B., Haigh, M., Harris, J., Overdahl, J., and Robe, M., 2008: Fundamentals, trader activity and derivative pricing. In *EFA 2009 Bergen Meetings Paper*.

[21] Büyüksahin, B., Haigh, M., and Robe, M., 2008: Commodities and equities: A "market of one"? *US Commodity Futures Trading Commission*. Working paper.

[22] Büyüksahin, B., and Harris, J., 2011: Do speculators drive crude oil futures prices? *Energy Journal*, **32**(2), 167–202.

[23] Büyükşahin, B., and Robe, M. A., 2014: Speculators, commodities and cross-market linkages. *Journal of International Money and Finance*, **42**, 38–70.

[24] Byun, S., 2014: The usefulness of cross-sectional dispersion for forecasting aggregate stock price volatility. *University of California, San Diego Working paper*.

[25] Byun, S., and Jo, S., 2014: Heterogeneity in the dynamic effects of uncertainty on investment. *University of California, San Diego Working paper*.

[26] Campbell, J. Y., Lettau, M., Malkiel, B. G., and Xu, Y., 2001: Have individual stocks become more volatile? an empirical exploration of idiosyncratic risk. *Journal of Finance*, **56**(1), 1–43.

[27] Carpenter, R. E., Fazzari, S. M., and Petersen, B. C., 1998: Financing constraints and inventory investment: A comparative study with high-frequency panel data. *Review of Economics and Statistics*, **80**(4), 513–519.

[28] Chirinko, R. S., 1993: Business fixed investment spending: Modeling strategies, empirical results, and policy implications. *Journal of Economic Literature*, **31**(4), 1875–1911.

[29] Connor, G., Korajczyk, R. A., and Linton, O., 2006: The common and specific components of dynamic volatility. *Journal of Econometrics*, **132**(1), 231–255.

[30] Considine, T., 1997: Inventories under joint production: An empirical analysis of petroleum refining. *Review of Economics and Statistics*, **79**(3), 493–502.

[31] Day, T. E., and Lewis, C. M., 1992: Stock market volatility and the information content of stock index options. *Journal of Econometrics*, **52**(1), 267–287.

[32] Deaton, A., and Laroque, G., 1992: On the behaviour of commodity prices. *Review of Economic Studies*, **59**(1), 1–23.

[33] Denis, D. J., and Sibilkov, V., 2010: Financial constraints, investment, and the value of cash holdings. *Review of Financial Studies*, **23**(1), 247–269.

[34] Diebold, F. X., and Mariano, R. S., 1995: Comparing predictive accuracy. *Journal of Business and Economic Statistics*, **13**, 253–263.

[35] Diebold, F. X., and Nerlove, M., 1989: The dynamics of exchange rate volatility: a multivariate latent factor arch model. *Journal of Applied Econometrics*, **4**(1), 1–21.

[36] Dixit, A. K., and Pindyck, R. S., 1994: *Investment under uncertainty*. Princeton university press.

[37] Duffie, D., 2001: *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey.

[38] Dunis, C. L., Laws, J., and Chauvin, S., 2001: The use of market data and model combination to improve forecast accuracy. *Developments in Forecast Combination and Portfolio Choice*, 45–80.

[39] Eckstein, Z., and Eichenbaum, M., 1985: Inventories and quantity-constrained equilibria in regulated markets: The us petroleum industry, 1947-1972. *Energy, foresight, and strategy.*

[40] Einloth, J., 2009: Speculation and recent volatility in the price of oil. *Federal Deposit Insurance Corporation.* Working paper.

[41] Eisner, R., 1978: Capital expenditure - some further analysis. In *Factors in Business Investment*, 117–132. NBER.

[42] Engle, R., and Sheppard, K., 2008: Evaluating the specification of covariance models for large portfolios. *New York University, working paper.*

[43] Engle, R. F., 1982: Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

[44] Engle, R. F., Ng, V. K., and Rothschild, M., 1990: Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics*, **45**(1), 213–237.

[45] Erb, C., and Harvey, C., 2006: The strategic and tactical value of commodity futures. *Financial Analysts Journal*, 69–97.

[46] Fama, E., and French, K., 1987: Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *Journal of Business*, 55–73.

[47] Fama, E., and French, K., 1988: Business cycles and the behavior of metals prices. *Journal of Finance*, **43**(5), 1075–1093.

[48] Fattouh, B., Kilian, L., and Mahadeva, L., 2012: The role of speculation in oil markets: What have we learned so far?

[49] Fattouh, B., and Mahadeva, L., 2012: Assessing the financialization hypothesis. *Oxford Institute for Energy Studies.* Working paper.

[50] Faulkender, M., and Wang, R., 2006: Corporate financial policy and the value of cash. *Journal of Finance*, **61**(4), 1957–1990.

[51] Franses, P. H., and Van Dijk, D., 1996: Forecasting stock market volatility using (nonlinear) garch models. *Journal of Forecasting*, 229–235.

[52] Gertler, M., and Gilchrist, S., 1994: Monetary policy, business cycles, and the behavior of small manufacturing firms. *Quarterly Journal of Economics*, **109**(2), 309–340.

[53] Ghosal, V., and Loungani, P., 2000: The differential impact of uncertainty on investment in small and large businesses. *Review of Economics and Statistics*, **82**(2), 338–343.

[54] Giacomini, R., and White, H., 2006: Tests of conditional predictive ability. *Econometrica*, **74**(6), 1545–1578.

[55] Gibson, R., and Schwartz, E., 1990: Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 959–976.

[56] Gilbert, C., 2010: Speculative influences on commodity futures prices 2006-2008. *University of Trento*. Working paper.

[57] Gilchrist, S., and Himmelberg, C., 1999: Investment: fundamentals and finance. In *NBER Macroeconomics Annual 1998, volume 13*, 223–274. MIT Press.

[58] Gorton, G., and Rouwenhorst, K. G., 2006: Facts and fantasies about commodity markets. *Financial Analysts Journal*, **62**(2).

[59] Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G., 2012: The fundamentals of commodity futures returns. *Review of Finance*, rfs019.

[60] Hamilton, J., 2009: Causes and consequences of the oil shock of 2007–08. *Brookings Papers on Economic Activity*, 215–284.

[61] Hamilton, J., 2009: Understanding crude oil prices. *Energy Journal*, **30**(2).

[62] Hamilton, J. D., 1994: *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.

[63] Hamilton, J. D., and Wu, J. C., 2014: Risk premia in crude oil futures prices. *Journal of International Money and Finance*, **42**, 9–37.

[64] Hansen, P. R., and Lunde, A., 2006: Consistent ranking of volatility models. *Journal of Econometrics*, **131**(1), 97–121.

[65] Hartman, R., 1972: The effects of price and cost uncertainty on investment. *Journal of Economic Theory*, **5**(2), 258–266.

[66] Hayashi, F., 1982: Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica: Journal of the Econometric Society*, 213–224.

[67] Herskovic, B., Kelly, B., Lustig, H., and Van Nieuwerburgh, S., 2014: The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Chicago Booth Research Paper*, (12–54).

[68] Hubbard, R. G., 1997: Capital-market imperfections and investment. *NBER Working Paper No. 5996*.

[69] Hwang, S., and Satchell, S. E., 2005: Garch model with cross-sectional volatility: Garchx models. *Applied Financial Economics*, **15**(3), 203–216.

[70] Irwin, S., and Sanders, D., 2010: The impact of index and swap funds on commodity futures markets. *OECD Food, Agriculture and Fisheries Working Papers*, **27**, 16–21.

[71] Jones, C. S., 2001: Extracting factors from heteroskedastic asset returns. *Journal of Financial economics*, **62**(2), 293–325.

[72] Kaldor, N., 1939: Speculation and economic stability. *Review of Economic Studies*, **7**(1), 1–27.

[73] Kilian, L., 2009: Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review*, **99**(3), 1053–1069.

[74] Kilian, L., and Murphy, D. P., 2014: The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics*, **29**(3), 454–478.

[75] Kilian, L., and Vigfusson, R. J., 2011: Nonlinearities in the oil price-output relationship. *Macroeconomic Dynamics*, **15**(S3), 337–363.

[76] Kydland, F., and Prescott, E., 1982: Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, 1345–1370.

[77] Kydland, F., and Prescott, E., 1988: The workweek of capital and its cyclical implications. *Journal of Monetary Economics*, **21**(2-3), 343–360.

[78] Lang, L., Ofek, E., and Stulz, R. M., 1996: Leverage, investment, and firm growth. *Journal of Financial Economics*, **40**(1), 3–29.

[79] Leahy, J., and Whited, T., 1996: The effect of uncertainty on investment: Some stylized facts. *Journal of Money, Credit and Banking*, **28**(1), 64–83.

[80] Merino, A., and Ortiz, A., 2005: Explaining the so-called 'price premium' in oil markets. *OPEC review*, **29**(2), 133–152.

[81] Ng, V., and Pirrong, S., 1994: Fundamentals and volatility: Storage, spreads, and the dynamics of metals prices. *Journal of Business*, 203–230.

[82] Pagan, A. R., and Schwert, G. W., 1990: Alternative models for conditional stock volatility. *Journal of Econometrics*, **45**(1), 267–290.

[83] Patton, A. J., 2011: Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, **160**(1), 246–256.

[84] Pindyck, R., 1994: Inventories and the short-run dynamics of commodity prices. *Rand Journal of Economics*, 141–159.

[85] Pindyck, R., 2001: The dynamics of commodity spot and futures markets: A primer. *Energy Journal*, **22**(3), 1–30.

[86] Pinkowitz, L., Stulz, R., and Williamson, R., 2006: Does the contribution of corporate cash holdings and dividends to firm value depend on governance? a cross-country analysis. *Journal of Finance*, **61**(6), 2725–2751.

[87] Pirrong, S., 2011: *Commodity Price Dynamics: A Structural Approach.* Cambridge University Press, Cambridge.

[88] Poon, S.-H., and Granger, C., 2005: Practical issues in forecasting volatility. *Financial Analysts Journal*, **61**(1), 45–56.

[89] Poon, S.-H., and Granger, C. W. J., 2003: Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, **41**(2), 478–539.

[90] Ramey, V., 1989: Inventories as factors of production and economic fluctuations. *American Economic Review*, 338–354.

[91] Routledge, B., Seppi, D., and Spatt, C., 2000: Equilibrium forward curves for commodities. *Journal of Finance*, **55**(3), 1297–1338.

[92] Sanders, D., and Irwin, S., 2010: A speculative bubble in commodity futures prices? cross-sectional evidence. *Agricultural Economics*, **41**(1), 25–32.

[93] Sarkar, S., 2000: On the investment–uncertainty relationship in a real options model. *Journal of Economic Dynamics and Control*, **24**(2), 219–225.

[94] Schwartz, E., and Smith, J., 2000: Short-term variations and long-term dynamics in commodity prices. *Management Science*, 893–911.

[95] Schwert, G. W., and Seguin, P. J., 1990: Heteroskedasticity in stock returns. *Journal of Finance*, **45**(4), 1129–1155.

[96] Singleton, K. J., 2013: Investor flows and the 2008 boom/bust in oil prices. *Management Science*, **60**(2), 300–318.

[97] Slade, M. E., 2013: Investment and uncertainty with time to build: Evidence from u.s. copper mining. *UBC working paper*.

[98] Stock, J. H., and Watson, M. W., 2002: Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, **97**(460), 1167–1179.

[99] Stoll, H., and Whaley, R., 2010: Commodity index investing and commodity futures prices. *Journal of Applied Finance*, **20**(1), 7–46.

[100] Tang, K., and Xiong, W., 2012: Index investment and financialization of commodities. *Financial Analyst Journal*, **68**(6), 54–74.

[101] West, K. D., 1996: Asymptotic inference about predictive ability. *Econometrica*, 1067–1084.

[102] West, K. D., and Cho, D., 1995: The predictive ability of several models of exchange rate volatility. *Journal of Econometrics*, **69**(2), 367–391.

[103] Weyant, J., and Huntington, H., 2010: Inventory of research and analysis on overall oil price determinants. *Energy Modeling Forum*.

[104] Working, H., 1949: The theory of price of storage. *American Economic Review*, **39**(6), 1254–1262.

[105] Ye, M., Zyren, J., and Shore, J., 2002: Forecasting crude oil spot price using oecd petroleum inventory levels. *International Advances in Economic Research*, **8**(4), 324–333.